Sec. 8

Journal of the AMERICAN STATISTICAL ASSOCIATION

UNIVERSITY OF MICHIGAN
JUN 22 1953
BUSINESS ADMINISTRATION LIBRARY

JUNE 1953

Wesley Clair Mitchell: The Economic Scientist ..... *Adolf A. Berle* 169

The Velocity of Time Deposits ..... *George Garvy* 176

Changes in the Functional Distribution of Income ..... *Jesse Burkhead* 192

The Post-Enumeration Survey of the 1950 Census: A Case History in Survey Design ..... *Eli S. Marks, W. Parker Mauldin, and Harold Nisselson* 220

On the Distinction Between Enumerative and Analytic Surveys ..... *W. Edwards Deming* 244

Confidence Intervals for the Number Showing a Certain Characteristic in a Population when Sampling is Without Replacement ..... *Leo Katz* 256

The Up-and-Down Method with Small Samples ..... *K. A. Brownlee, J. L. Hodges, Jr., and Murray Rosenblatt* 262

Designing Single-Sampling Inspection Plans when the Sample Size is Fixed ..... *Abraham Golub* 278

On Errors in Matrix Inversion ..... *Paul S. Dwyer and Frederick V. Waugh* 289

A Multiple Group Least Squares' Problem and the Significance of the Associated Orthogonal Polynomials ..... *Bradford F. Kimball* 320

Estimation of the Parameters of Type III Populations from Truncated Samples ..... *Des Raj* 336

Proceedings of the 112th Annual Meeting 350

BOOK REVIEWS by Colin Clark, C. C. Craig, Churchill Eisenhart, Harry P. Hartkemeier, E. L. Lehmann, G. J. Lieberman, A. W. Marshall, Daniel O. Price, Norman Rudy, Paul B. Simpson, R. Clay Sprowls. 358

PUBLICATIONS RECEIVED 382

RANDOM DIGITS 383

VOLUME 48 NUMBER 262

Now every typist can be a statistical typist. Nine keys, built right into the keyboard of an IBM Electric Typewriter, take the drudgery out of tabular typing.

The mere touching of the proper decimal tabular key speeds the carriage automatically to the correct amount position within the column. No backspacing, no waste hand motion, no lost time.

Write today for a demonstration or illustrated brochure on the IBM Decimal Tabulation model to: IBM, Dept. JS, 590 Madison Avenue, New York 22, N. Y.

IBM *Electric Typewriters* INTERNATIONAL BUSINESS MACHINES

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 48 JUNE 1953 NUMBER 262

ARTICLES

Wesley Clair Mitchell: The Economic Scientist . . ADOLF A. BERLE 169
The Velocity of Time Deposits . . . . . GEORGE GARVY 176
Changes in the Functional Distribution of Income. . JESSE BURKHEAD 192
The Post-Enumeration Survey of the 1950 Census: A Case History in Survey Design . . . . . ELI S. MARKS, W. PARKER MAULDIN, AND HAROLD NISSELSON 220
On the Distinction between Enumerative and Analytic Surveys . . . . . W. EDWARDS DEMING 244
Confidence Intervals for the Number Showing a Certain Characteristic in a Population when Sampling is without Replacement . LEO KATZ 256
The Up-and-down Method with Small Samples . . . . . K. A. BROWNLEE, J. L. HODGES, JR., AND MURRAY ROSENBLATT 262
Designing Single-sampling Inspection Plans when the Sample Size is Fixed. . . . . ABRAHAM GOLUB 278
On Errors in Matrix Inversion . . . . . PAUL S. DWYER AND FREDERICK V. WAUGH 289
A Multiple Group Least Squares' Problem and the Significance of the Associated Orthogonal Polynomials . . . . BRADFORD F. KIMBALL 320
Estimation of the Parameters of Type III Populations from Truncated Samples . . . . . DES RAJ 336
Proceedings of the 112th Annual Meeting 350

BOOK REVIEWS

WILSON, E. BRIGHT, *An Introduction to Scientific Research* . . . . . E. L. LEHMANN 358
ANDERSON, R. L., AND BANCROFT, T. A., *Statistical Theory in Research* . . . . . CHURCHILL EISENHART 359
HAGOOD, MARGARET JARMAN, AND PRICE, DANIEL O., *Statistics for Sociologists, Revised Edition* . . . . . A. W. MARSHALL 362
GRANT, EUGENE L., *Statistical Quality Control, Second Edition* . . . . . C. C. CRAIG 363
HALD, A., *Statistical Theory with Engineering Applications* . . . . . G. J. LIEBERMAN 365
FREUND, JOHN E., *Modern Elementary Statistics* . . . NORMAN RUDY 367
BLAIR, MORRIS MYERS, *Elementary Statistics, Revised* . . . . . R. CLAY SPROWLS 368
MCMILLEN, WAYNE, *Statistical Methods for Social Workers* . . . . . DANIEL O. PRICE 370
CASEY, ROBERT S., AND PERRY, JAMES W., editors, *Punched Cards: Their Applications to Science and Industry* . HARRY P. HARTKEMEIER 372
BENNETT, MERRILL K., *Population, Food, and Economic Progress* . . . . . COLIN CLARK 374
BUREAU OF THE CENSUS AND BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM, *Census of Manufactures: 1947: Indexes of Production* . . . . . PAUL B. SIMPSON 376
ROBINSON, MARILYN DRUCK, *Washington State Statistical Abstract* . . . . . PAUL B. SIMPSON 380
PUBLICATIONS RECEIVED. 382
RANDOM DIGITS 383

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

The Editors welcome the submission of manuscripts for possible publication. They should be typewritten entirely double-spaced, including footnotes, and two copies should be sent to the Editor, W. Allen Wallis, 207 Haskell Hall, University of Chicago, Chicago 37. Books for review should be sent to the same address. Unsolicited book reviews are not accepted, but suggestions of titles for review are welcome.

EDITOR

W. ALLEN WALLIS, *University of Chicago*
ASSISTANT TO THE EDITOR: MARGARET A. LABADIE

ASSOCIATE EDITORS

HOWARD L. JONES
Illinois Bell Telephone Co.
GEORGE M. KUZNETS
University of California
WILLIAM G. MADOW
University of Illinois
PHILIP J. MCCARTHY
Cornell University
I. RICHARD SAVAGE
National Bureau of Standards
C. ASHLEY WRIGHT
Standard Oil Company (N.J.)

ADVISORY PANEL OF FORMER EDITORS

WILLIAM G. COCHRAN (1945–50)
Johns Hopkins University
WILLIAM F. OSBORN (1920–1925)
University of Chicago
FRANK A. ROSS (1926–34, 41–45)
Thetford, Vermont
FREDERICK F. STEPHAN (1935–40)
Princeton University

Errata: Readers and authors are urged to submit to the Editor notices of errors found in this or any previous issue. These will be published once a year, in the December issue.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 282 JUNE 1953 *Volume 48*

WESLEY CLAIR MITCHELL: THE ECONOMIC SCIENTIST*

ADOLF A. BERLE, JR.
Columbia University

THE legacy of a scholar is his teaching. The tribute to a scholar is the esteem of his peers. The immortality of a scholar is the influence his thinking exerts on generations yet to come. Wesley Clair Mitchell can fairly be named the greatest American economic scholar of the twentieth century. Laying a foundation for the evolution of scientific economics was his precise and monumental contribution. Why it was done at all and how it was achieved is a subject of first importance. Because of Mitchell's work, economics can fairly claim now to be called a "science," just as, for lack of a similar contribution in political science and sociology, these sisters of economics still grope in the gray fog that lies between the area of speculation and the explored territory of generalization and conclusion from observed facts.

Colleagues, disciples and students bear affectionate and discerning witness in this volume to his achievement. It is fitting that they have done so because of their relationship with the man; but, more important, it is useful that they have done so because the book is a case-study of the process by which economics was translated from a kind of theology into a more or less exact science. The essays here collected by Arthur F. Burns thus are far more than a memorial; they are a contribution to the study of economic methodology, and cannot fail to be causative.

Before Mitchell tackled the field of economic method, economics consisted of speculative development of hypotheses formulated by earlier, intelligent observers of their environment. These attempted generalization from their limited and personal observation. Such was the work of Adam Smith, of Ricardo, of John Stuart Mill, and later of Marshall and of Taussig. Their speculations opened a wide door (of

* A review of *Wesley Clair Mitchell; The Economic Scientist*, edited by Arthur F. Burns, (New York: National Bureau of Economic Research, Inc., 1952).

this, more later). Without them, no science could ever have been constructed; but acceptance of their hypotheses without demonstrating their factual base clearly left economics in the realm of inductive philosophy rather than of scientific demonstration. It remained for Mitchell to pioneer the job of isolating identifiable economic factors, determining analytic definitions of phenomena, working out methods of accurate measurement and continuing observation, verifying relationships or at least establishing probability of a relationship of cause and effect between phenomena. Classification, description, measurement, establishment of interrelationships, generalization, looking toward prediction, this in economics as elsewhere is the basis of scientific method. As a young man, Mitchell was unsatisfied by speculation without proof, and he accordingly introduced scientific method into economics with modesty and integrity, combined with a good humor which made him at once the most beloved as well as the most respected of contemporary social scientists.

How he did it is partly a saga of life, partly a demonstration of close thinking. In one sense, he never transcended an early conception blocked out when he was studying at the University of Chicago. There he was working on that grimmest of assignments, a Ph.D. thesis, designed to be "A History of the Legal Tender Acts." But it evolved into "History of the Greenbacks," and became eventually a powerful analysis of monetary theory in the United States. His intellectual godfathers in that period were surely diverse: Thorstein Veblen, brilliant, tormented, biting; John Dewey, challenging conventional process of thought; Jacques Loeb, psychologist, and physiologist, devoted without limit to scientific method. This constellation was something for students to dream of: a brilliant analytic mind, a sceptical but human philosophical mind, and a rigorous scientific methodological mind. Mitchell absorbed wisdom from all three; and an economist was formed. In unpredictable result, the concept of a true science of economics was really generated through Mitchell by a tempestuous dreamer, a revolutionary philosopher and a pessimistic physiologist.

The consequence of this amazing combination of stimuli (carefully analyzed in an essay by Professor Frederick C. Mills) was Mitchell's application of methods known in physical science in the form of a massive monograph on "Business Cycles." It was magnificently and accurately factual—a description rather than an explanation. At the close of his life, Mitchell was still working on business cycles. He still considered that he had not solved the problem. But as a result, the "business cycle" had been identified if not explained. The National

Bureau of Economic Research, in whose organization Mitchell had played the leading part, had made notable progress in the vast task of analyzing, describing, and collecting accurate data in sequence. Measurement and actual calculations of "national income" and of "gross national product," the analysis of the rise and fall of price levels, all stem from this life-time work. "The notion that inquiries should be framed from the start in such a way as to permit of testing the hypothetical conclusions," the notion of sequence, the concept of consecutive growth as opposed to older ideas of assumed but unexplained (and actually non-existent) equilibrium, are perhaps the most causative of Mitchell's ideas. Hundreds of men are working now on lines of investigation opened up by them.

The bibliography of his work, beginning with "The Quantity Theory of the Value of Money" (1896) and closing with "What Happens during Business Cycles: A Progress Report" (1951), is a catalogue of hundreds of studies, large and small, of exact, factual data, so handled that each piece of measurement or description could be related to and used in connection with other collections of fact-data. Inspired by his influence, a host of men have since followed his example. At the close of his life, great areas of economic theory have been or can be tested by reference to relevant facts. A demonstrated conclusion can be distinguished from unverified theory. Economics has moved steadily forward; it is fairly on its way.

The essays here collected show all this, though each naturally follows the turn of its respective writer. Mills, perhaps chiefest heir of Mitchell's thought, places his professional sketch next to a tender and penetrating memoir written by Lucy Sprague Mitchell. The two are not as far apart as might be supposed. Research which tests dogma against fact is always controversial and in some measure impersonally destructive and it may even be revolutionary. An honest man may find himself intellectually bound to refute by his facts the most tenaciously held doctrine, and is not popular when he does so. A true scientific scholar must set his lance above mischance and ride the barrière. The professional disciple and the wife find much common ground in their descriptions.

Mitchell was always working on two levels: thus, in 1920 he was founding the National Bureau of Economic Research with the impersonal scientific end of measuring magnitude of national income and measuring its principal components. But in the same year he joined James Harvey Robinson and Alvin Johnson in founding "The New School for Social Research" as a part of his life-long protest against

restricting freedom of speech in academic institutions and somewhat later he was endeavoring to salvage the remaining years of Thorstein Veblen from the dark limbo into which Veblen's stormy soul had steered him. Mitchell gave the best of his thought to governmental research and planning, which necessarily exacts a high degree of organized conformity. But to the end of his life he worked at experimental education—which is anything but conformist. To say of him, as does John Maurice Clark in his reprinted Memorial address, "American economics has lost its undisputed first citizen," is a plain statement of fact.

Method, well discussed by A. B. Wolfe, is worth a word. We shall always have with us the academic individual who fears to make a hypothesis or propound a theory and confines himself to the less controversial task of measuring phenomena. But the fact is that measuring phenomena without *some* hypothesis to determine what facts to collect is sterile business. Even a fact-collector must have a hypothesis—else a count of paving stones in Times Square would be just as significant as calculating the size and distribution of the national debt. Mitchell never fell into the error of first accumulating facts and then constructing a hypothesis. He knew better than most that the hypothesis guides the collection of facts; but that from the collected facts, the hypothesis can be tested, revised, or discarded and new hypotheses can be made. On the other hand, he never committed to any body of theory, and thereby provoked a major debate in certain quarters.

In the one critical essay included in the volume, Paul T. Homan observes that, as a sequel of Mitchell's method, increasingly sound scientific work is being devoted to detailed studies, disclosing the manner of the workings of portions of the economic system "to the ultimate end of assisting in intelligent social guidance. Realism is Allah, and Mitchell is his prophet. Paradise may be around the corner." Unless, suggests Homan, you commit to some sort of theory, you are lost. Implying that Mitchell did not commit himself, he ironically adds that there is no scientific reason to suppose that the "voluntary process of cumulative causation is amenable to intelligent social control. . . . There is no guide here but faith." Which is another way of saying that value judgment of Mitchell's scientific economics depends on philosophical premises; that Mitchell had one but did not state it and could not test it; that for all his realist detachment, Mitchell's work was really energized by quite homely, non-scientific convictions, and that he might as well have stated them as theory.

The implicit criticism is, of course, that without a comprehensive

theory which Mitchell never embraced and never formulated, the significance of his work can not be appraised. This criticism is fair enough though not too damaging. Actually, as a result of Mitchell, any economic theory, formulated or hereafter formulated, will be held valid only if verified by observable data, and testing for some time to come will probably be by use of the statistical tools Mitchell forged or by methods he developed. Development of method was Mitchell's chosen function; fundamentally he was a master toolmaker. One suspects he also cherished (by inheritance from Veblen) a hidden and unfulfilled desire to write a great Utopia, but consciously limited himself to statistical work lest he become intoxicated with his own dreams. But—men are led by their dreams and their dreamers, as politicians know all too well. Statistical analysts can merely record the results of their pursuit, in hope that succeeding generations may perhaps be better guided in selecting what manner of dreams and dreamers to follow.

If Mitchell's name is attached to anything, it will be to a school of economic thought now known as "institutionalism." In Wolfe's words, "Economics must be a concrete and realistic study of institutional habits and relations, and these must be regarded not as fixed but as continually changing; hence economics is not only an institutional but a dynamic and evolutionary science." To this is linked another aphorism that economics has only one justification—the furtherance of economic welfare. Men, *en masse*, behave in certain standard ways, according to uniform though complex patterns. These patterns are in constant evolution, but in part at least they can be measured by statistics. Patterns of human action tend to be reflected in institutions or persistent groupings; they include, for instance, the American corporation (the connection in which this reviewer first met Wesley Mitchell) or the habits and framework of consumer credit or possibly a Communist commissariat or twenty other similar institutional operations. These institutional patterns become, in themselves, factors of greater or less force in economic action. All this transcends what was commonly known as "orthodox" economic theory. (Milton Friedman makes this point with discernment in his chapter on "The Economic Theorist.") It accounts for Mitchell's frank recognition that the money economy (commonly miscalled "capitalist" economy) is by no means the only possible system in which economic principles have validity, or in which economic science can exist.

* * *

Because the chief authors of these essays are economists, any lay reviewer must submit his comments with diffidence. No criticism of

Mitchell is implied in the observation that his insistence on demonstrated facts isolated him from some of the driving conflicts of his time. Facts are few and hard come by; to insist on them before asserting a viewpoint does accomplish detachment. Yet detachment is, and in Mitchell's case was, bought at a price, whose extent is just beginning to be visible. If, for example, the dominant motive activating human patterns ceases to be individual search for profit, and becomes rather a reaching for power, individually or through collective groups, much of current economic theory would have to be revised and statistical conclusions would assume different significance. Only in 1952 an economist's first attempt at reappraising the American capitalist system in this sense has reached print in Kenneth Galbraith's *American Capitalism: The Concept of Power*. The thesis would not be undreamed of in Mitchell's philosophy—but it would be excluded from his economic work.

Again, certain bench marks sacred to economists, such as the "basic interest rate," may be a result of inevitable tides; or they may be the result of more or less consciously managed monetary systems. Individual search for security and comfort, indeed, may perhaps give way under some circumstances to a mass desire, or at least willingness, to be dedicated or subjected to directed collectivism, as for example, in the Soviet Union. It is quite conceivable that as modern technology progresses, the power motive rather than the profit motive may become the key to institutional development. It is at least possible that the titanic world struggle of today arises from the tensions created by some such shift. It could be, in a word, that economics apart from political science is meaningless abstraction. Detachment from political struggle may not be a luxury permitted to economists in days to come. Mitchell never ignored such struggles; but he never mixed his thoughts about them with his economic work.

In like manner, Mitchell's insistence that measures looking toward social reorganization must be based on "established knowledge" is merely Utopian. Statesmanship usually has to meet crises arising out of the unknown or unexplained. Historians may later supply the "established knowledge" needed for full diagnosis; but the statesman, businessman, or politician rarely has the luxury of awaiting that happy day. He must act in accordance with someone's best estimate of probabilities, and the event will prove him right or wrong. For an economist to refuse to estimate in these circumstances is scientifically creditable, but it deprives the politician of the putative best estimate.

* * *

For laymen, much of the economist's language is almost unintelligible. But they understood Mitchell. Most educated people today have some idea of the national income of the United States and of its fluctuations, and the national government uses his statistical instrument as a chief factor in determining fiscal policy. Laymen can and do make use of the growing body of data covering its use and its distribution. Businessmen as well as economists have a gradually growing knowledge of the relation of money and credit to the rise and fall of employment and trade, and to the creation and distribution of goods and services. Thousands of investors who never heard of Mitchell know about the business cycle, and seek to anticipate it in conducting their affairs. Hundreds of bankers have unostentatiously changed their financial theory because of his critique of the quantity theory of money. Impact of his ideas is found all over the American governmental and business system; it is no light recognition of his accomplishment.

Yet even today Mitchell's greatest premise has not been fully understood. Nearly thirty years ago he indicated it:

> In becoming consciously a science of human behavior, economics will lay less stress upon wealth and more stress upon welfare. Welfare will mean not merely abundant supply of serviceable goods but also a satisfactory working life filled with interesting activities. . . .

He dreamed of developing "criteria of welfare" as well as of statistical measures for income and wealth. In his view, this science of economics, struggling to be born, was but one tool in the greater search for life—however conceived.

Mitchell's great contemporary, Professor Joseph Schumpeter of Harvard, a few days before he also joined Mitchell in the Valhalla of scholars, wrote in closing the record:

> Here was a man who had the courage to say, unlike the rest of us, that he had not all the answers; who went about his task without either haste or rest; who did not care to march along with flags and brass bands; who was full of sympathy with mankind's fate, yet kept aloof from the market place; who taught us, by example and not by phrase, what a scholar should be.

THE VELOCITY OF TIME DEPOSITS*

GEORGE GARVY
Federal Reserve Bank of New York

THE rate of turnover of time deposits belongs to the outer reaches of the *terra incognita* of the behavior of liquid assets. "Liquid assets" and not "money" is intentionally used because the question whether time deposits are part of the money supply is one on which there are differences of opinion among monetary theorists.[1] Even within the Federal Reserve System there is no unanimity: while the Board of Governors includes in the money supply not only time deposits of commercial banks, but also deposits of mutual savings banks and of the Postal Savings System, the Federal Reserve Bank of New York does not include any of these items.[2]

Several writers on monetary problems have inquired into the relative velocities of circulation of various components of the money supply, but only after making their own decisions on the proper definition of money.[3] Their explorations, moreover, generally involve certain assumptions as to the velocity (and other dimensions) of one or several components of the total money supply for which empirical data are not readily available. Such assumptions with respect to time deposits range from Keynes' view that savings deposits have a "velocity of zero" to Burgess' more generally used estimate of a rate of turnover of twice a year.[4] Most recent writers either confine themselves to the statement that the rate of turnover of time deposits at commercial banks is low or very low,[5] or fall back on Burgess' estimate.

It is not intended to review here the controversy on the "proper" definition of the money supply. Indeed, a meeting of minds is more likely to be advanced by an inquiry into the behavior of time deposits

---

* This paper is based, in part, on a report prepared for a technical committee of the Federal Reserve System.

[1] See, for instance, A. Marget's exhaustive discussion, in *The Theory of Prices*, Vol. I, Ch. XVI.

[2] See, for instance, the article "Money Supply" in the *Monthly Review of Credit and Business Conditions* of the Federal Reserve Bank of New York, Nov. 1951.

[3] See Marget, *op. cit.*, Vol. I, p. 463, footnote 10.

[4] W. R. Burgess, "Velocity of bank deposits," *Journal of the American Statistical Association*, 1923, pp. 727–40. Burgess' estimate, however, was based on fragmentary data obtained in 1922 from only six New York City banks for short periods, usually one month in each case. The validity of his generalization of this rate for the country as a whole may be challenged on the basis of data which have become available subsequently and which show that time deposits in other parts of the country turn over less rapidly than in New York City.

[5] See, for instance, L. Currie, *The Supply and Control of Money in the United States* (Harvard University Press, 1934), p. 50. Angell does not inquire into the turnover of time deposits because "there appears to be no data for measuring their rate of turnover in the past." See *The Behavior of Money* (McGraw-Hill Book Co., 1936), p. 94.

than by further exploration of the logical possibilities of defining attributes of money. If it is agreed that some types of liquid assets, such as currency and demand deposits, are money *par excellence*, the approach might be to inquire whether time deposits (and other classes of liquid assets) show similar velocity and similar cyclical and seasonal patterns.[6] If business firms and consumers consider time deposits as equivalent to money, the behavior of time deposits should bear a close resemblance to that of currency and demand deposits. For the present study, the behavior of demand deposits will serve as a convenient frame of reference since comprehensive data on the velocity and patterns of use of currency are lacking.

Because of the paucity of data for commercial bank time deposits, the scope of this inquiry has been broadened to include savings institutions. While the great bulk of deposits of savings institutions consists of savings deposits of individuals, time deposits of commercial banks[7] also include time deposits of their trust departments, of business firms, and of government units. Such deposits are frequently held in the form of certificates of deposit or in open accounts. The shift in the composition of time deposits accounts, as will be shown below, for the divergent movement of the rate of turnover of time deposits and of savings deposits.

Data on the turnover of *total time deposits* at commercial banks are available for only a short period prior to World War II. They have been collected from about 400 "money market" banks (weekly reporting member banks) in New York City and 100 other leading cities for the relatively short period from September 5, 1934 to February 1, 1939. Data on the velocity of *savings* deposits only, were collected, on an annual basis only, by the Savings Division of the American Bankers' Association from about 140 scattered commercial banks for the period 1940–47; similar rates for 1949–51 were obtained from more than 800

[6] It should be clear from the above that the concern here is with the "moneyness" of time deposits rather than with other aspects of the problem, such as the fact that time deposits provide reserves for credit expansion or that it is difficult to separate the two types of deposits. These two points have been used by Goldenweiser (in his recent *American Monetary Policy*) as arguments for including time deposits with the money supply.

The fluidity of the concept of money has been increased by the war-time growth of liquid assets possessing a high degree of convertibility and attaining annual turnover rates similar—as a matter of fact, identical—to those of time deposits. Thus, Series E Savings bonds, after being held for at least sixty days, now are as easily and conveniently converted into cash as time deposits. Furthermore, while legally a 30-day notice can be enforced against holders of savings accounts, Savings bonds are payable on demand. Obviously, all types of liquid assets possess some degree of "moneyness." Marget's argument that *all* deposits have *some* rate of turnover—be it only once every 1,000 years—may be extended to include all types of liquid assets.

[7] U. S. Government and interbank time deposits are always excluded from the term "time deposits" as used in this paper.

commercial banks. These data can be supplemented by time series for the turnover of deposits at mutual savings banks, savings and loan associations, and the Postal Savings System which can be computed for periods of various lengths. In addition, several special surveys conducted by the Savings Banks Trust Company help to shed additional light on the behavior of savings deposits.

THE RELATIVE IMPORTANCE OF SAVINGS DEPOSITS IN COMMERCIAL BANK TIME DEPOSITS

The justification for drawing limited inference from the behavior of savings deposits to the behavior of total time deposits in commercial banks lies in the fact that the share of the other types of time deposits in total time deposits is known to have been small and declining. The earliest available data on the proportion of savings deposits in time deposits are limited to national banks, fewer than half of which reported savings accounts separately for June 30, 1910. In these banks, savings deposits accounted for about 57 per cent of a derived total, "time deposits";[8] by June 1921, the percentage of savings deposits and open accounts combined (no separate figures being available) was over 73 per cent, and by June 30, 1926, over 79 per cent. Time certificates of deposit, which in 1917 accounted for nearly 40 per cent of all time deposits, in 1926 represented only 21 and in 1929 16.5 per cent.

In subsequent years, the importance of certificates of deposit and of open accounts declined not only relatively, but also absolutely, as shown in the following table for *all member banks* of the Federal Reserve System:

Call date	Savings deposits	Time certificates of deposit	Open accounts*	Total time deposits
	(In millions of dollars)			
December 31, 1928	9,810	1,895	1,089	12,794
June 30, 1942 (latest date available)	10,357	566	749	11,673

* Including small amounts in Christmas savings and similar accounts.
Source: Board of Governors of the Federal Reserve System, *Member Bank Call Reports.*

[8] Prior to the Federal Reserve Act, national banking legislation did not specifically distinguish between time and demand deposits. Since the classification of deposits reported to the Comptroller of the Currency was left entirely to the judgment of the reporting banks (see *Annual Report of the Comptroller of the Currency* for 1910, p. 7), and because of incomplete coverage, the percentage for 1910 is hardly more than an approximation.

After 1921, the classification "savings deposits" had become identical with the subtotal "time deposits other than certificates" and was reported by all national banks having such deposits.

All time deposits, used here, exclude State and municipal and Postal Savings deposits.

By June 30, 1945 (the latest date for which such data are available for all national banks), the percentage that savings deposits represented of all time deposits of individuals, partnerships, and corporations at *national* banks had risen to 95.9 per cent. It was slightly lower—94.0 per cent—for all commercial banks largely because of the inclusion of State trust companies.

Although no national data are available for the postwar years, those for State commercial banks and trust companies in New York State indicate that the prewar decline of time deposits other than savings deposits is still continuing. Even though the accelerated growth of the relative importance of savings deposits during the war years was reversed in the first postwar years, at the end of 1951 savings deposits represented a higher percentage of total time deposits than at the end of 1940. This was true for banks located outside New York City (where the ratio in earlier years was close to the national average) as well as for those in New York City (where it was much lower than for banks outside New York City because of the importance of trust department and other types of open accounts):

Savings Deposits as a Percentage of Total Time Deposits of Individuals, Partnerships, and Corporations

End of	New York City	Rest of New York State
1940	41.7	95.4
1945	65.0	98.1
1951	47.7	97.1

Source: New York State Banking Department.

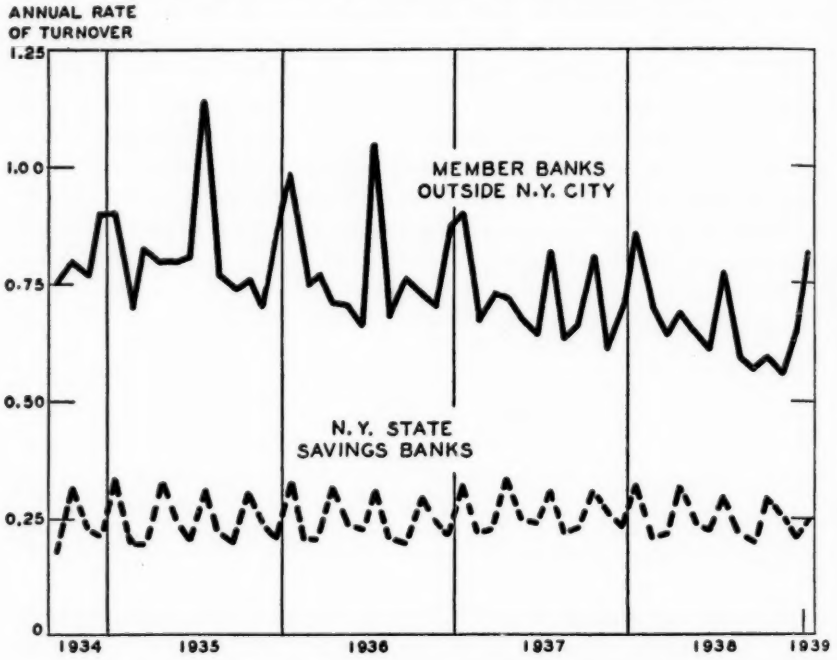
Thus, the long-run decline of the relative importance of time deposits other than savings deposits is apparently still continuing. Outside New York City (and possibly outside Chicago and one or two other money centers) time deposits other than savings accounts are, however, already reduced to such a small proportion of the total that their further decline is unlikely to have much effect on the velocity of total time deposits.

TURNOVER RATES OF TIME AND SAVINGS DEPOSITS

During the years immediately preceding World War II, *time* deposits in weekly reporting member banks outside New York City turned over from 0.66 to 0.82 times a year (Annual averages; for monthly data, see Chart I). However, time deposits at commercial banks include some

deposits of business firms, local governments, and trust departments, which tend to be more active than savings deposits. Since the weekly reporting banks include essentially money market banks with a relatively high proportion of such deposits, it is likely that the velocity of

CHART I. Annual Rates of Turnover of Time Deposits* in Weekly Reporting Member Banks Outside New York City, and of Regular Deposits† in New York State Mutual Savings Banks. September 1934 through January 1939.



* Total time deposits except interbank between September 1934 and January 1938; thereafter time deposits of individuals, corporations, etc., States and political subdivisions. See footnote 6, page 177.

† Including school savings accounts.

Source: Board of Governors of Federal Reserve System, and Savings Banks Trust Co.

time deposits at *all* commercial banks outside New York City was *lower* than the rates computed from figures reported by these money market banks.

For the same reason, the velocity of time deposits in New York City (where, in recent years, savings deposits may have accounted for as

little as 50 per cent of time deposits) ranged between 1.65 and 1.99.[9] In comparison, demand deposits at the weekly reporting member banks in New York City turned over about forty times as rapidly as time deposits.[10]

A detailed analysis of these data was made at the Board of Governors of the Federal Reserve System only for the single year 1936, when the annual rates of turnover ranged between 0.5 (Philadelphia District and New York District outside New York City) and 1.1 (San Francisco District). In New York City, time deposits were more active, but still turned over less than twice (1.4 times) a year as compared with 0.8 time for all 100 outside centers. The higher turnover rates in New York City were traced to relatively active time deposits of trust departments of some banks and in the San Francisco District to a large proportion of rather active local government funds. It was also suggested that reporting (contrary to instructions) of renewals of maturing certificates of deposit as debits had likely resulted in overstating the velocity of time deposits, especially in the Boston and Minneapolis Districts (where certificates of deposit are more important than in other districts). The reporting of transfers of time deposits to demand deposits is likely to have had the same effect.[11]

The results of the ABA surveys are consistent with the rates based on Federal Reserve System studies. They show that annual rates of turnover of *savings* deposits at cooperating commercial banks ranged in 1940–51 between 0.45 and 0.60;[12] the rapid growth of savings deposits which have more than doubled during this period may be regarded as a sufficient explanation for the slight downward drift in the average rate of turnover over the twelve-year period.

---

[9] The rates of turnover of time deposits in weekly reporting New York City banks reflect primarily activity of time accounts other than personal savings accounts. Their monthly fluctuations during the years 1935–39 were large and did not display the seasonal pattern shown by outside reporting centers or by savings banks. Increases in the rate of time deposit turnover of 50 or more per cent from one month to another were not infrequent. Such fluctuations most likely reflect unusual withdrawals from corporate, municipal government, and trust department time deposits.

While the rate of turnover of time deposits in banks outside New York City drifted downward during the period under review, the corresponding rate for New York banks rose.

[10] On the turnover of demand and of total deposits, see George Garvy, *The Development of Bank Debits and Clearings and Their Use in Economic Analysis* (Washington, 1952), Ch. VII.

[11] Turnover rates for individual banks were studied in the Boston District only. Wide variations were found, partly reflecting differences in reporting and classification practices of banks.

[12] A careful inspection of rates of turnover for individual banks in the ABA sample for 1940–47 reveals that in most banks, these rates were within the range of 0.3 and 0.6, which may be considered narrow in view of the differences in the structure and mobility of population, banking facilities, accounting practices, age of the savings departments, and other pertinent factors.

While institutions with very large savings deposits generally have higher turnover rates than those which hold a smaller volume, there are exceptions. As a matter of fact, some banks with upward of $100,000,000 of savings deposits ranked among those with the lowest turnover rates.

*Annual Rate of Turnover of Savings Deposits Savings Departments in Commercial Banks**

Year		Year	
1940	0.53	1946	0.60
1941	0.58	1947	0.56
1942	0.54	1948	—
1943	0.49	1949	0.45
1944	0.47	1950	0.48
1945	0.49	1951	0.46

* The survey for 1940–1947, conducted in 1948, covered 140 banks; the survey for 1949–51, made in 1952, included 808 banks.

Source: American Bankers' Association.

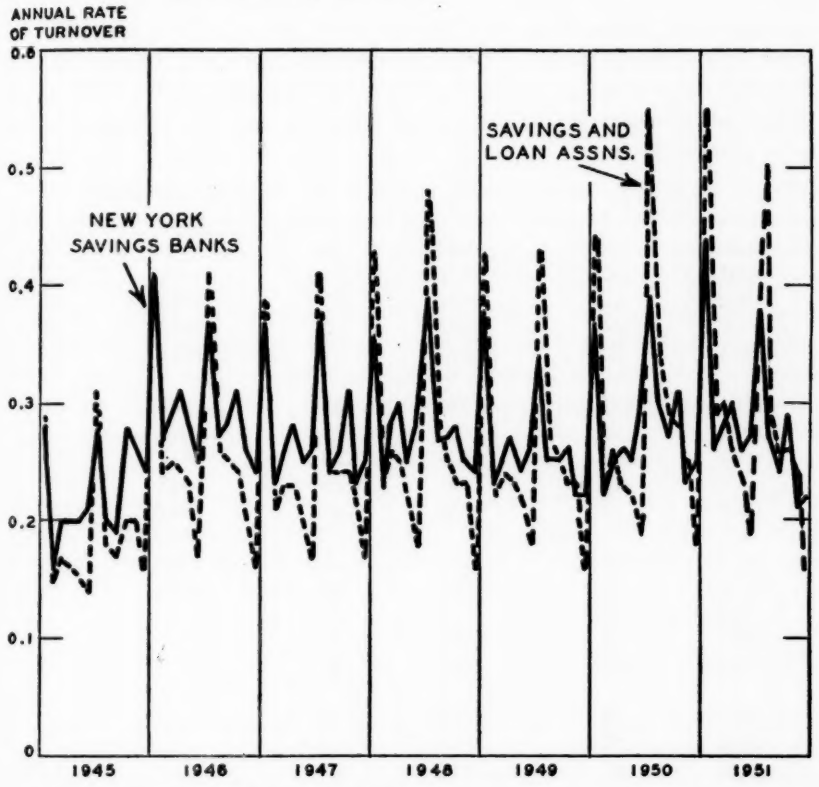
In contrast, deposits of mutual savings banks have been turning over since the end of World War II, on the average, once every four years; rates of turnover of shares of savings and loan associations since 1945 have been of the same magnitude (see Chart II). These velocity ratios are about half as large as those of savings deposits at commercial banks according to the ABA surveys. The principal reason for these differences in turnover rates lies in the broader coverage of commercial bank time deposits, and in institutional differences between the several types of savings institutions. They are, however, not of sufficient magnitude to preclude drawing some inference as to the behavior of turnover rates of time deposits prior to 1934 from data for mutual savings banks. Such data can be computed for New York State institutions for more than half a century.[13]

Since 1900, the annual rate of turnover of New York State mutual savings bank deposits has fluctuated within the relatively narrow range of 0.24 and 0.34, with the exception of 1929 (when it was 0.40) and most war years (when it was somewhat lower). During the twenties, it fluctuated at a higher average level than during the first two decades of the century or the years since 1934. In recent years, the velocity of savings deposits in New York savings banks has been still considerably lower than in the twenties, although somewhat higher than during World War II years. It reflects clearly the speculative peak of 1929 (when the rate of turnover rose to a level of one-third above that of the preceding years), the contractions in 1907 and 1937, and the uncertainties caused by the declarations of war in 1917 and 1941. The increase

[13] Turnover rates for New York State savings deposits (which are heavily weighted by New York City institutions with higher than average deposit activity) are somewhat higher than those for all U. S. mutual savings banks, but they exhibit month-to-month fluctuations which are very similar to those of the U. S. series. About half of the total national deposits is accounted for by New York State.

in 1950–51, traceable to the outbreak of hostilities in Korea, was relatively minor.

CHART II. Annual Rates of Turnover of Regular Deposits* in New York State Mutual Savings Banks and of Private Share Capital in all U. S. Savings and Loan Associations. Monthly 1945–51.



* Regular deposits including school savings accounts.

Source: For savings and loan association data, Home Loan Bank Board. For mutual savings bank data, Savings Banks Trust Co.

Turnover rates for New York State savings banks suggest that an increase in the public's desire to hold currency, resulting from an impairment of public confidence or from other causes,[14] is reflected in withdrawals of savings deposits as well as of demand deposits. The velocity of savings deposits, however, rises very little or not at all in periods of

[14] For a detailed analysis based on investigations made by several New York savings banks, see Weldon Welfling, *Savings Banking in New York State* (Duke University Press, 1939), pp. 156 ff. Dr. Welfling remarks, however: "The stock market activity would not necessarily produce an increase in withdrawals and accounts closed, for an increase in loans on passbooks would serve the same purpose to the depositor."

business expansion. Thus, the rate of turnover of deposits of New York savings banks did not increase in the late 1920's until 1929, lagging behind the velocity of demand deposits at commercial banks outside New York City which rose nearly one-third during this period. Again, while velocity of demand deposits increased by one-third from the low point in 1949 to the peak in the spring of 1951, the rate of turnover of savings bank deposits (and of shares of savings and loan associations) rose very little, perhaps 10 per cent. As already noted, it declined during the expansion of business activity during World War II, when, however, the activity of demand deposits at weekly reporting banks also slowed down.

Rates of Turnover of Deposits in New York State Savings Banks

1900	.26	1920	.33	1940	.24
1901	.26	1921	.32	1941	.27
1902	.25	1922	.31	1942	.25
1903	.26	1923	.33	1943	.22
1904	.26	1924	.32	1944	.21
1905	.27	1925	.33	1945	.23
1906	.29	1926	.32	1946	.30
1907	.31	1927	.32	1947	.28
1908	.28	1928	.33	1948	.29
1909	.25	1929	.40	1949	.27
1910	.26	1930	.32	1950	.29
1911	.26	1931	.34	1951	.29
1912	.26	1932	.32		
1913	.26	1933	.32		
1914	.25	1934	.26		
1915	.24	1935	.26		
1916	.22	1936	.26		
1917	.26	1937	.27		
1918	.26	1938	.25		
1919	.30	1939	.24		

Source: Savings Bank Trust Company. *Chart Book.*

Similarly, the rate of turnover of savings deposits in the ABA sample of commercial banks *declined* during the period of expanding business activity, 1941–44, and again after the reconversion, 1947 through 1949.[15]

[15] The lack of responsiveness of the rates of turnover of savings deposits to an expansion in business activity is confirmed by the behavior of Postal Savings deposits. The annual rate of turnover of Postal Savings *declined* from 0.89 in 1921 to 0.68 in 1925 and 0.61 in 1928. Similarly, it dropped from 0.70 in 1940 to 0.47 in 1945.

The annual rate of turnover of such deposits has been two or three times higher than that for New York State savings deposits, except for the first years following the institution of the Postal Savings System in 1911, when the rates were even higher because deposits were small.

Slight increases occurred in 1945 and 1950 and a rather sharp rise in the reconversion year 1946.

The relative stability of the rate of turnover of New York State *savings* deposits over a period of more than fifty years is entirely compatible with a declining trend in the rate of turnover of *total time deposits*,[16] which reflects the gradually diminishing importance of time deposits other than those evidenced by passbooks. There are two additional reasons for presuming that, prior to the Banking Act of 1933, time deposits at commercial banks were more active than during the period covered by the Board's study of weekly reporting member banks:

1. In the 1920's, it was possible to make withdrawals from savings accounts in some banks without personal presentation of passbooks. Opinions differ regarding the extent to which holders of time deposits considered them practically equivalent to demand deposits and were actually able to draw checks against them.[17] There is a voluminous literature on the subject, and views range from Schumpeter's assertion that "in the twenties time and demand deposits were essentially the same kind of thing"[18] to Currie's reluctance to consider time deposits essentially different from holdings of Government securities.[19]

It is uncertain whether the possibility of facile drawing against time deposit accounts was actually of material effect on the rate of withdrawals. On the other hand, after 1933, the prohibition of the payment of interest on demand deposits and the introduction of service charges (and their subsequent increases) may have induced some small depositors to make more active use of their savings accounts to cash checks and to purchase cashier's checks for making payments rather than to maintain separate but onerous checking accounts.

2. It is also frequently claimed that, before 1933, part of the time deposits in commercial banks actually represented misclassified demand deposits.[20] There are no direct quantitative data, and none are likely to become available, to resolve the conflicting views on the meaning of the growth of time deposits at commercial banks during the

---

[16] From the end of 1934 to 1939, when the turnover rates of the more homogeneous New York savings deposits fluctuated around a horizontal level, the velocity of time deposits at weekly reporting member banks outside New York City clearly showed a downward drift (Chart I).

[17] See Federal Reserve System *Report of the Committee on Member Bank Reserves* (Washington, 1931), p. 7.

[18] J. A. Schumpeter, *Business Cycles* (New York, 1939), Vol. II, p. 856.

[19] Currie, *op. cit.*, p. 14.

[20] See, for instance, B. M. Anderson's comments in *The Chase Economic Bulletin* (Nov. 8, 1926), p. 14 n. As already mentioned, this view was also held very strongly by Schumpeter. Woodlief Thomas wrote: "In the 1920's, they [time deposits] could be easily withdrawn from many banks, and sometimes included funds that might otherwise have been in demand deposits." (*Banking Studies*, p. 302).

twenties. The statistical evidence (supplementing personal knowledge of the existing practices) used by those who championed the opposing views, seems to have been limited essentially to data on the rate of growth of time deposits in different classes of banks and on the average size of time deposit accounts. The growth of time deposits, rather than any direct evidence of the classification of active deposits in the time category, gave rise to the views held by Anderson and others that during the twenties part of time deposits actually represented demand obligations.[21]

It is impossible to arrive at a firm conclusion as to the extent to which bankers might have misclassified demand deposits or given special drawing privileges to owners of time deposit accounts. In any case, in recent years more rigid standards applied by bank examiners must have made the use of time deposits for current transaction purposes extremely difficult (even though increased reserve requirements for demand deposits have made misclassification of deposits more attractive to bankers). It seems reasonable to assume that the velocity of time deposits has been declining over time primarily because of the decreasing influence of the more active time certificates of deposit and open accounts. Since these types of accounts are now reduced to insignificance except in large money centers, it is likely that the velocity of time deposits will exhibit a stability similar to that of savings bank deposits.

INTERPRETATION OF THE VELOCITY OF TIME DEPOSITS

The concept of velocity of deposits has a precise analytical meaning only if withdrawals can be equated with use for transactions purposes. It is, however, not any more feasible to distinguish statistically in the case of time than in the case of demand deposits between payments for goods, services, or the acquisition of assets and the mere transfer of balances between different depositaries or accounts of the same holder.

Such fragmentary data as are available seem to indicate that activity of time and savings accounts represents transfers of balances

---

[21] Professor French, who has carefully examined the statistical data for national banks for 1922–28 concluded that "the bank statistics show clearly that there could have been no appreciable amount of shifting of deposits from the demand to the time category on the part of large depositors in national banks. If there was any such shifting, it was on the part of small depositors. We may safely say that the same conclusion would be valid for other commercial banks." See D. R. French, "The Significance of Time Deposits in the Expansion of Bank Credit, 1922–28," *Journal of Political Economy*, Dec. 1931. French's statistical analysis was confirmed by his interviews with executives of California corporations holding time deposits.

between savings institutions,[22] shifts among alternative forms of investment or accounting entrances (such as expiration and renewal of certificates of deposit) at least as much as it represents payments for goods and services.[23]

To clarify the significance of the turnover rate of savings deposits, two aspects of account activity are of particular interest: (1) the seasonal pattern of turnover rates (which is nearly identical with that of withdrawals, since savings balances are very stable) and (2) the close month-to-month correspondence between the amount of new deposits and withdrawals.

Most New York savings banks pay interest at the end of June and December, while others credit interest quarterly. The seasonal pattern of turnover of deposits in New York savings banks (based on 1928–45) shows four peaks following these quarterly interest dates.[24] The highest peak occurs in January when rates of turnover are about 50 per cent higher than in the two immediately preceding and two following months; the July peak is somewhat lower, while the April and October activity is only about 30 per cent higher than in the eight months of slow activity.[25]

This pattern, which reflects that of withdrawals from savings accounts, may be explained largely in terms of interest payment dates.[26] Since the general practice is to credit interest only for amounts remain-

[22] This situation is indeed paralleled when demand deposits are shifted among several accounts (at the same or at different institutions) of the same individual or firm. Since, however, the transaction velocity of demand deposits is so much higher, the proportion of transfers to total transactions is likely to be considerably lower than in the case of time deposits.

[23] See Weldon Welfling, "Some Characteristics of Savings Deposits," *American Economic Review*, December 1940, pp. 748–58.

[24] Quarterly data on withdrawals collected by the National Bureau of Economic Research (for one New York savings bank for 1880–1923 and for five such banks for 1900–1947) consistently show peaks in the first and third quarters, corresponding to the January and July peaks of the more inclusive but shorter series used here.

[25] The available data for time deposits of weekly reporting member banks outside New York City show a broadly similar seasonal pattern, but the April and October peaks are much less pronounced than those in January and July.

The seasonal pattern of withdrawals from savings and loan associations (based on 1944–51 data for the entire United States) shows two sharp peaks, in January and July (which are even sharper than for New York mutual savings banks) reflecting payment of dividends which in some cases is made automatically by check. The rates of turnover of the private share capital of U. S. savings and loan associations decline gradually during each of the five months following the two months of peak withdrawals (if the small dip in February is removed by placing the index on the daily average base).

In contrast, December is the peak month for rates of turnover of demand deposits at the weekly reporting member banks, although the predominance of December has been declining in the moving seasonal factors computed by the Federal Reserve Bank of New York. November is the second highest month of demand deposit velocity. January is an average (or slightly above average) month, while July has been consistently below average, reflecting the slowing down of business activity during the vacation period. High rates of turnover of demand deposits in the two last months of the year reflect Christmas business, the movement of crops, and a multiplicity of business and personal end-of-year payments.

[26] Tax payment dates and seasonal need for extra cash (for payment of Christmas charge accounts in January and for vacation money in July) might be additional influences.

ing on deposit during the entire interest period, depositors who do not need funds urgently usually prefer to postpone withdrawals until the end of the interest period in order to earn the full interest. In particular investors who transfer their accounts from one institution to another would do it normally after interest for the last perod has been credited. Individuals, such as retired people, who use the income from savings accounts (and the principal as well) to meet current expenditures, are likely to make withdrawals after the (semiannual or quarterly) interest payment dates. Finally, since there is in New York State an upper limit (prior to 1951, 7,500 dollars) on individual savings bank accounts, interest on deposits which are at the maximum is normally withdrawn because no interest is paid on the excess amounts. Although the proportion of the accounts which are at the legal limit is small, they account for a relatively large share of total deposits.

While interest payment dates provide a sufficient explanation for the peaks in withdrawals (and in the rate of turnover) we must seek other reasons for the very close correspondence between monthly withdrawals and new deposits. Indeed, as Chart III shows, high monthly withdrawals are usually associated with high levels of new deposits and low withdrawals with low deposits.

There are several possible explanations for such a close correlation. A transfer of funds from one savings bank to another is reported both as a withdrawal and a deposit. A more important explanation seems to be, however, that a large proportion of account activity represents use of savings accounts for the purpose of cashing occasional checks; many people deposit checks to their savings accounts and withdraw part or the entire amount in cash.[27]

Thus it is likely that, in addition to interest payments, the January peak in account activity reflects cashing of dividend, bonus, and Christmas checks. In other quarterly months, similar deposits of dividend and other quarterly checks are made and a large proportion of the proceeds is withdrawn in cash.

The conclusion seems justified that velocity rates for *savings* deposits

[27] There are among savings accounts a number of small but relatively active accounts which are apparently maintained primarily to cash checks rather than to build up a backlog of savings. An analysis of a sample of widely scattered savings departments of commercial banks and mutual savings banks, revealed that in 1937–40 most accounts which averaged more than one withdrawal a month accounted for less than 5 per cent, and in only exceptional cases for as much as 10 or more per cent, of total deposit liability. Such accounts were responsible, however, for from 10 to 50 per cent of the number of withdrawals in individual institutions.

A much smaller sample showed that in 1943–44 the proportion of active accounts declined, but that such accounts were considerably more active than during the years immediately preceding World War II. Most likely, these changes reflect temporary wartime developments (see Savings Banks Trust Company, *Chart Book*, Chart A-26).

CHART III. Monthly Deposits* and Drafts as Per Cent of Opening Amount Due Depositors Each Month 1928–1946. All Savings Banks in New York State.

are not necessarily indicative of the rate at which *balances* are used for transaction purposes. Part of the activity of savings deposit accounts reflects shifts of investments among various media, including transfers from one bank to another. Rates of savings deposit turnover reflect the use of balances, but also the use of the product of such balances (interest) and—perhaps even more—the utilization of check cashing facilities which these balances provide.[28]

Inter-city differences in rates of *time* deposit turnover also reflect differences in the relative importance and activity of time deposits of business firms and of government units and in accounting practices. All this, together with the existence of a strong seasonal pattern, makes the interpretation of changes in the rates of time deposit turnover rather difficult.

The main conclusions of this paper may be summarized as follows:

1. Savings deposits at commercial banks turn over about once every two years, or only about one-fiftieth as rapidly as demand deposits.

2. Because time deposits (other than interbank and U. S. Government) at commercial banks include several types of time accounts in addition to savings deposits, their velocity is higher than that of savings deposits alone, but still less than one. (The relatively higher turnover of time deposits in New York City member banks compared with "outside" banks reflects a higher percentage of time deposits other than savings deposits.)

3. There has been a downward drift of the rate of time deposit turnover in commercial banks outside New York City reflecting the relative increase of savings deposits in total time deposits of these institutions. The relative proportion of savings deposits in reported time deposits has been increasing continuously since the establishment of the Federal Reserve System, thus narrowing the difference between rates of turnover of commercial bank time deposits and savings deposits. Currently, more than 94 per cent of time deposits at commercial banks outside New York City are savings deposits. It is, therefore, likely that *changes* in the velocity of savings bank accounts approximate closely those of savings and also time deposits at commercial banks outside New York City.

4. Deposits at savings banks and shares of savings and loan associations turn over about half as rapidly as savings deposits at commercial

---

[28] A survey made of several savings institutions during the last war indicated that checks cashed or partly cashed (a portion of the cash being deposited) were between ¼ and ½ the number of regular deposit and draft transactions. (See Savings Banks Trust Company, *Chart Book*, Chart A-26.) In some cases, checks are chased in full as a courtesy, without requiring them to be deposited.

banks. This difference in *level* of velocity (changes being similar) is attributable to institutional factors.

5. The rate of turnover of savings deposits is more stable than that of demand deposits. This stability reflects partly the failure of savings deposits to become more active when business expands.

6. The difference in the nature and use made of savings and demand deposits is reflected in significant differences in the seasonal patterns of their velocities.

7. A relatively small proportion of savings accounts with nominal balances are used for cashing salary checks and for the temporary safekeeping of funds causing them to influence disproportionately the activity of total deposits of a savings institution.

If these conclusions based on fragmentary and heterogeneous data surveyed have general validity, it would appear that those who are reluctant to include time deposits in the money supply can find in them substantial arguments for their position. Indeed, if distinction between money and other types of liquid assets is approached not as a logical but as an empirical problem, it is apparent that average turnover rates of between 20 and 50 times a year are significantly different from those of once every two or four years. A significant part of the "activity" of savings accounts does not reflect withdrawals for making payments, but arises from the closing of accounts, from the cashing of salary and other checks (which are ultimately debited against demand deposit accounts at commercial banks thus affecting velocity data based on such debits), and from the transfer of funds (including interest received) to other savings institutions. Analysis of the available statistical data on the rates of turnover of savings and time deposits shows, furthermore, that these rates do not exhibit the same cyclical conformity or the same seasonal patterns as the rates of turnover of demand deposits.

Considerably more justification could be found for including time deposits other than savings deposits with the money supply, although for recent years the amounts involved are very small in relation to demand deposits.

CHANGES IN THE FUNCTIONAL DISTRIBUTION OF INCOME

JESSE BURKHEAD*
Syracuse University

THE increased quantity and quality of national income statistics which have become available within the past few years make possible a re-examination of significant relationships within the U. S. economy. The present study represents an examination of one such relationship—labor and property shares in the national income. For this purpose the detailed accounts prepared by the National Income Division of the Department of Commerce have been utilized for the years 1929–1950. The conclusions for this period are then compared with the Kuznets data for the years 1919–1929.

The present study is intended to serve two purposes:

First, it will examine the common generalization that wages and salaries are a stable proportion of national income.[1] This generalization turns out to be valid only in relation to a particular set of national income concepts. When other concepts are employed it appears that labor shares in national income are not stable, but shift significantly from year to year and over a period of years. Section I will be devoted to an examination of alternative national income concepts which will differ substantially from those employed by the Department of Commerce. Data on relative shares (changes in the distribution of income by type) will be presented and examined in Section II and in Section III.

Second, this study will investigate changes in labor and property shares in the last three decades, but with particular reference to the decade of the Forties. That is, attention will be centered on the behavior of labor and property shares under inflationary conditions. The data show a general pattern: increases in economic activity are associated with a reduction in labor shares of total income; decreases in economic activity are associated with an increase in labor shares of total income; inflationary conditions appear to accentuate the decline in labor shares which is associated with increases in activity. The rationale for this behavior pattern will be considered in Section IV.[2]

---

* The author is indebted to Joseph A. Pechman of the U. S. Treasury Department for criticisms of an earlier draft of this article.

[1] See, for example, Paul A. Samuelson, *Economics* (McGraw-Hill Book Co., 1951), p. 231.

[2] In the literature, the problem of inflation has generally been treated in aggregative terms, as an "excess of demand over supply," in terms of the rate of inflation and its measurement as a departure from assumed stability conditions. See, for example, Arthur Smithies, "The behavior of money national income under inflationary conditions," *Quarterly Journal of Economics*, November 1942, 113–138; Tjalling Koopmans, "The dynamics of inflation," *Review of Economics and Statistics*, May 1942, 53–65. See also the literature on the inflationary gap, such as Walter A. Salant, "The inflationary gap," *American Economic Review*, June 1942, 308–313; Milton Friedman, "Discussion of the inflationary gap," *ibid.*, 314–20.

Before examining the data in detail it will be necessary to set forth the conceptual basis for the measurement of labor and property shares in total income.

I

National income concepts have no independent validity; they must be appropriate to the purposes at hand.[3] The present purpose is the measurement of the relative magnitude of labor income and of property income over a period of time. Two basic concepts (and some variants thereof) will be employed. The first basic concept will be labeled the "primary" distribution of income. The second, the "claims" concept of income.

The primary distribution of income will measure labor and property shares derived from current economic activity. Therefore, income equals earnings paid or accrued for the use of property or personal services. This concept is similar to "earning power" measurements of income;[4] and to the concept of producers' income.[5]

The second basic concept—income claims—attempts to measure labor and property shares in terms of relative economic power. Claims on goods and services reflect not only earnings and accruals from current output but also earnings from the purchase and sale of assets (net capital gains) and from income receipts which are not derived from current production. The claims concept of income is, therefore, equal to the primary distribution of income plus net capital gains plus transfer payments. This brings the concept closer to Haig's definition of income as "the money value of the net accretion to one's economic power between two points of time."[6]

Much less attention has been paid to structural problems—to the changes *within* the economy during inflationary periods. In one of the few contributions to this subject Holsman points out that only two of the ten economists represented in the symposium on inflation in the August, 1949 *Review of Economics and Statistics* mentioned distributional aspects of inflation (See Franklyn D. Holsman, "Income determination in open inflation," *Review of Economics and Statistics*, May 1950, 150–58). Another contribution to structural relationships is Ralph Turvey, "Period analysis and inflation," *Economica*, August 1949, 218–27.

[3] Simon Kuznets, "National Income," *Encyclopedia of the Social Sciences* (New York, 1934), Vol. XI, p. 206.

[4] For a discussion of the "earning power" concept of income see, *Concentration and Composition of Individual Incomes, 1918–1937*, TNEC Monograph No. 4, Washington, 1940 (by Adolph J. Goldenthal), pp. 9–11.

[5] Tibor Barna, *Redistribution of Incomes Through Public Finance in 1937* (Oxford, 1945), p. 17.

[6] R. M. Haig, *The Federal Income Tax* (Columbia University Press, 1921), p. 7. (Cited by Roy Blough and W. W. Hewett, "Capital gains in income theory and taxation policy," *Studies in Income and Wealth*, Vol. II (National Bureau of Economic Research, 1938), p. 197.) Haig would define "net accretion" to include consumer expenditures during the period. A complete application of the "economic power" concept of income would require the inclusion of both realized and unrealized net capital gains, but no data are available on the latter. Also, it would be necessary to adjust income receipts and accruals in accordance with personal taxes. No attempt has been made to deal with this adjustment on the data presented here. However, an effort along these lines, based on Department of Commerce income concepts, has been recently attempted. See Edward F. Denison, "Distribution of national income," *Survey of Current Business*, June 1952, 22–3.

The remainder of this section will be devoted to examination of the conceptual basis of the primary distribution of income, and the behavior of labor and property shares when measured on this basis. The claims concept of income will be examined in Section III.

PRIMARY DISTRIBUTION OF INCOME

The measurement of labor and property shares in accordance with the primary distribution of income is set forth in detail for the years 1929–1950 in Appendix Table A. Three concepts of total income have been employed. Labor income is identical in each of these three. Concepts II and III differ from Concept I in their treatment of property income.

Labor income (Concepts I, II, III) includes:

1. The compensation of employees

This embraces wages and salaries in the private, military and government civilian sectors, and the supplements to wages and salaries, including employer contributions to social insurance and to private pension funds. The compensation of corporate officers is treated here as labor income, although a part of such compensation should probably be classified as property income. That is, corporation managers may be in a position to "divert" corporate net income from stockholders to managers. The amount of such diversion is likely to be small in relation to the total of property income and will probably exert no perceptible influence on relative shares.[7]

2. Income of unincorporated enterprise less inventory valuation adjustment.

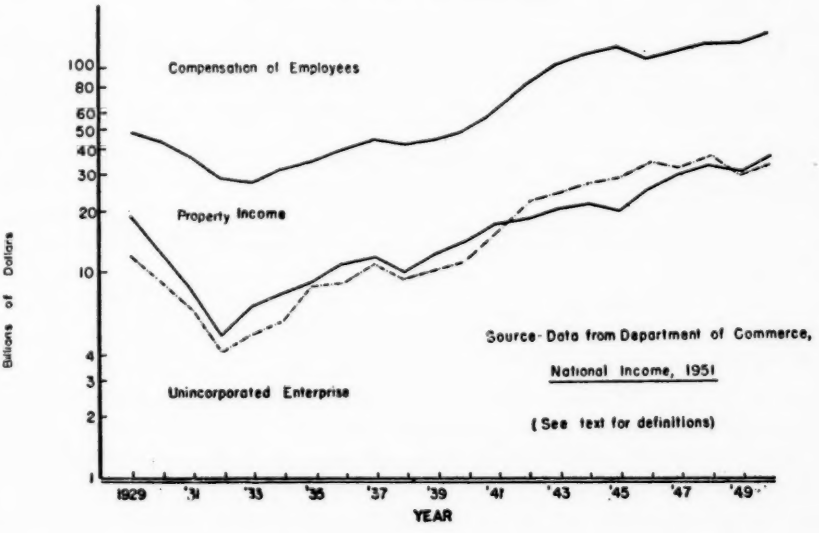
It is exceedingly difficult to determine whether this type of income is a labor or a property share. Entrepreneurial income is a mixture of payment for personal services (labor income) and return from property. In some instances, as with the income of professionals, it may be predominantly attributable to labor. In other instances, as in unincorporated manufacturing enterprise, it may be predominantly a return to capital. Not only will the division of this type of income between labor and property vary from industry to industry, it will also vary within an

[7] For example, in 1939 the compensation of corporate officers was about 9 per cent of the total of wages and salaries in the corporate sector. In 1948 it was about 7.5 per cent (U. S. Department of Commerce, *National Income, 1951*, p. 156). On the other hand, there has undoubtedly been a growth in pension plans, stock purchase plans and expense accounts available to executives. It is not possible to estimate the volume of this kind of income or determine whether it represents property income or labor income.

industry. Furthermore, there is no reason to assume that the labor-property relationship within the total of income from unincorporated enterprise will remain constant from year to year.

The problem of classification of income of unincorporated enterprise becomes, then, the logically impossible feat of separating the inseparable parts of a whole. However, the sector is too large to be ignored; in 1950 it accounted for 15 per cent of the total of national income. Chart I shows the movements in the totals of property income, labor income and income from unincorporated enterprise for the years 1929–1950. No clear pattern of relationship is established by the fluctuations in the three series. During the depression and recovery years the income of unincorporated enterprise generally moved in the same direction and to the same degree as property income. But in the war years its movements were comparable to those of employee compensation.

CHART I. Compensation of Employees, Property Income, Income of Unincorporated Enterprise, 1929–1950.



The industrial composition of income of unincorporated enterprise may provide a somewhat better basis for its classification as labor income or property income than its behavior pattern. Although the composition of income of unincorporated enterprise varies from year to year, four categories generally comprise its greatest bulk: agriculture, retail trade, medical and other health services, and legal services. (In

1950 these four made up 72 per cent of the income of the sector.) Since these four categories, in turn, appear to be predominantly labor-type activities, it would appear most reasonable to classify the whole sector as labor income. This would appear to be a better procedure than attempting to estimate its changing labor and property segments year by year.[8]

The classification of income from unincorporated enterprise as labor income would therefore appear to be justified on the ground of its industrial composition. The advantage of this classification is that it obliterates the shifts between the self-employed and wage earners; that is, it eliminates changes in labor shares that would otherwise result from changes in the proportion of the labor force which is self-employed. The inescapable disadvantage is that labor income, defined to include the income of unincorporated enterprise, will contain an uncertain and variable component of property income.

Property income, Concept I, includes:

1. Rental income of persons

Although this type of income reflects some element of personal services provided by landlords to tenants, no breakdown of this portion is possible and the receipt as a whole is predominantly property income.

2. Corporate profits after tax, without allowance for inventory valuation adjustment.

This measurement of property income in the corporate sector raises two questions:

a. In the consideration of income shares derived from current production it is appropriate to measure corporate profits after tax rather than corporate profits before tax. Although corporate profits before tax are, in some sense, available for the payment of wage increases, once the wage bill is paid, corporate profits *taxes* are not available to either labor or property income recipients; the taxes may not be spent by the corporation or its stockholders. The measurement of corporate profits net of taxes has the further advantage that it puts all property income on a uniform basis with respect to taxation, that is, all property income is then measured net of business taxes but gross of personal taxes.[9]

[8] Kuznets seems to be inclined toward the view that entrepreneurial income is predominantly labor income. See Simon Kuznets, *National Income and Its Composition, 1919–1938*, Vol. I (National Bureau of Economic Research, 1941), p. 82.

[9] The controversy concerning whether corporate profits taxes are reflected in product prices is not germane to the approach here, where income is measured in terms of shares paid or accrued.

However, for illustrative purposes, a measurement of property income which includes corporate profits taxes is set forth in Concept II. The differences in behavior exhibited by the two concepts will be examined below.

b. Inventory valuation adjustment is appropriately excluded from the measurement of property shares. The Department of Commerce adjustment is for the purpose of arriving at the value of goods and services produced within the accounting period. But profits or losses on inventory should be included in a measurement of payments or accruals to property income recipients from current activity. Such profits and losses are as much property income, and as much dependent on current operations, as profits from the production and sale of goods and services within the accounting period.

3. Net interest

This receipt is clearly and unambiguously property income.

4. Net government interest

Although excluded from national income, and treated, in effect, as a transfer payment in the Department of Commerce national income accounts, government interest is included here in the total of property income. Viewed from the standpoint of the recipient of government interest, this type of receipt is as much property income as interest on industrial bonds. The recipient of interest is the owner of property; this property yields a return in the form of government interest payments. The inclusion of government interest is, therefore, appropriate when income flows are viewed from the standpoint of the recipient. It is equally appropriate to exclude government interest when, as in the Department of Commerce concepts, income flows are viewed in relation to the aggregate of current output.[10]

The sum of items 1–4 represent property income, Concept I. Table 1 shows a reconciliation between the concept of national income employed by the Department of Commerce and the income concepts employed here for the measurement of the relative shares of labor and property income paid or accrued to labor and property recipients.

Two additional concepts of property income have been developed in

[10] In accordance with Department of Commerce procedure, government interest is measured exclusive of intra-governmental transfer, with the assumption that this amount may be appropriately imputed to individuals; that is, government interest receipts by financial institutions correspond to imputed service charges. (See *National Income, 1951*. p. 93.)

examining relative shares in accordance with the primary distribution of income. In Concept II corporate profits taxes have been added. This concept is developed primarily for purposes of comparison, and in accordance with the convention that corporate profits taxes should be reflected in the national income.

In Concept III property income reflects the addition of capital consumption allowances (corporate profits taxes are excluded). This treatment has the advantage of establishing a parity between labor and

TABLE I

RECONCILIATION: NATIONAL INCOME (DEPARTMENT OF COMMERCE), THE PRIMARY DISTRIBUTION OF INCOME, THE CLAIMS CONCEPT OF INCOME, 1947

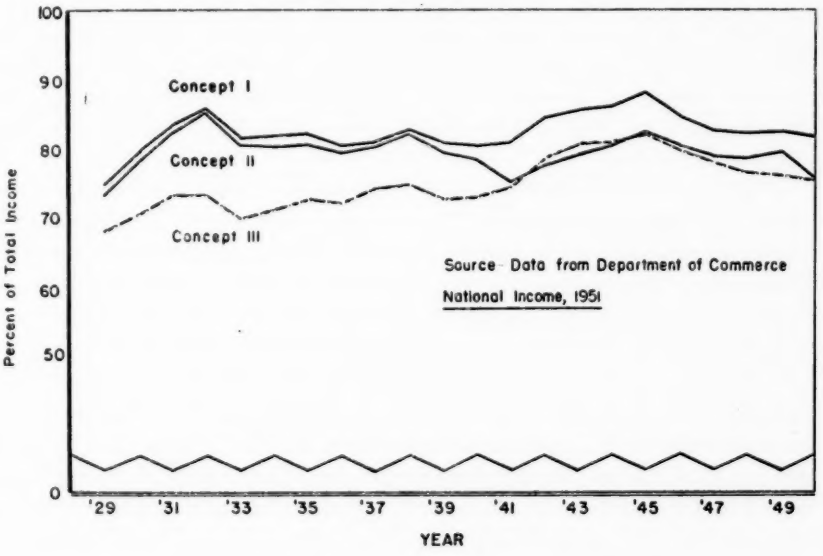
		(Millions of dollars)
National Income, Department of Commerce		198,688
Deduct:		
Inventory Valuation Adjustment, Corporate	−5,757	
Inventory Valuation Adjustment, Unincorp.	−1,547	
Corporate Profits Taxes	11,940	4,636
Add: Government Interest		4,378
The Primary Distribution of Income:		
Concept I		198,430
Add: Corporate Profits Taxes		11,940
Concept II		210,370
Concept I		198,430
Add: Capital Consumption Allowances		14,845
Concept III		213,275
The Claims Concept of Income:		
Concept I (primary distribution		198,430
Add: Net Capital Gains, Corporate		924
Net Capital Gains, Individuals		4,450
Transfer Payments		11,803
Deduct: Contributions for Social Insurance		7,268
Concept I		208,339
Add: Capital Consumption Allowances		14,845
Concept III		223,184

property shares, since labor income, of necessity, includes "labor consumption allowances," that is, the cost of maintaining labor's earning capacity. Capital consumption allowances, as estimated by the Department of Commerce, represent amounts which are generally available to the property sector for purposes of maintaining capital equipment. Whether or not these allowances are adequate to maintain the capital "intact" is irrelevant for present purposes. The important point is that capital consumption allowances are available for expenditure by the property sector.

II

The behavior of labor and property shares in total income, where total income is based on the primary distribution of income, is depicted in Chart II. The underlying data are shown in Appendix Table A. The three concepts which have been discussed are plotted in terms of relatives, that is, as labor's share of total income for each year 1929–1950.

CHART II. Labor Shares in Total Income: the Primary Distribution of Income, 1929–1950.



FLUCTUATIONS IN THE PRIMARY DISTRIBUTION OF INCOME

Labor shares, measured in accordance with Concept I, rise sharply as economic activity declines from 1929 to 1932. With the progress of recovery after 1933, labor shares decline at the outset and then remain

reasonably stable for the remainder of the decade at from 80 to 83 per cent of total income. This is a substantially higher share than in 1929 when labor shares were 75 per cent of total income. Since labor-property shares in 1929 were comparable to labor-property shares in most of the years of the Twenties, it follows that labor's share of total income in the late Thirties was substantially higher than in the Twenties.

The war period exhibits a striking increase in the share of total income going to labor, in marked contrast with the generally stable relationship from 1933 to 1941. The war economy distributed a substantially higher proportion of total income to labor than in the late Thirties. In 1945 labor's share of total income was 88 per cent. This was 7 percentage points more than in 1941 and 13 percentage points more than in 1929.

The postwar period, marked by a decline in real output and an increase in prices, is associated with a decline in labor's share of total income. The downward movement carries labor shares almost but not quite back to the levels of 1933–1941. Interestingly enough, the 1949 recession is marked by a slight increase in labor shares—consistent with the slight increase in the recession of 1938. Also there is a slight downturn in labor shares in 1950, notable not so much for the magnitude of change, which is insignificant, but for the direction of change. From 1941 to 1942 the mobilization program produced a sharp increase in labor's share of total income. This seems not to have happened in the 1950 mobilization.

The behavior of labor shares in total income, where income includes corporate profits taxes, is shown in Concept II, Chart II. The year to year movements are comparable to those of Concept I, where corporate profits taxes are excluded, except for the years 1939 to 1941 when corporate profits taxes rose sharply. However, the inter-decade position of labor income is different when Concept II is employed. In World War II labor shares of total income (including corporate profits taxes) now appear to be no higher than in the late Thirties, and, in the post-war period, substantially lower than in the late Thirties. As has been argued above, it would not be appropriate to attach too much importance to conclusions based on a concept of income which includes corporate profits taxes, since these taxes should not properly be viewed as income paid or accrued from current activity.[11]

[11] A recent study of wage shares in the national income of the United Kingdom shows a very different behavior pattern than for the United States. The war years, for example, show low rather than high ratios of wage income to the total. (See E. H. Phelps Brown and P. E. Hart, "The Share of Wages in National Income," *Economic Journal*, June 1952, 253–77, especially p. 265.) However, the income concepts employed by Messrs. Brown and Hart are not wholly comparable with those utilized here. Profits, for example, are measured gross of profits taxes.

In Concept III, Chart II, capital consumption allowances are included in property income, and a sharply different behavior pattern emerges than when capital consumption allowances are neglected. Property income shares are not only considerably larger but now show much greater stability through the years of most serious depression. In 1932 capital consumption allowances are larger than all other property income.

The inclusion of capital consumption allowances further accentuates the relative improvement in labor's share during World War II. The shrinkage in labor shares in the postwar period is similar in magnitude to that exhibited by Concept I. However, Concept III shows that labor's share remains higher in the postwar period than in the Thirties. This is to be attributed to the fact that capital consumption allowances are smaller in relation to other income in the postwar period than during the Thirties.

Perhaps the most important conclusion which emerges from the data shown in Chart II is that, employing either Concept I or Concept III, labor shares in total income rose sharply in World War II, receded during the postwar period, but remained somewhat higher than in the late Thirties. The economy appears to have been distributing a larger portion of its total income to labor over the past twenty years, although the postwar inflation has offset a part of this tendency.

PRIMARY DISTRIBUTION OF INCOME IN THE CORPORATE SECTOR

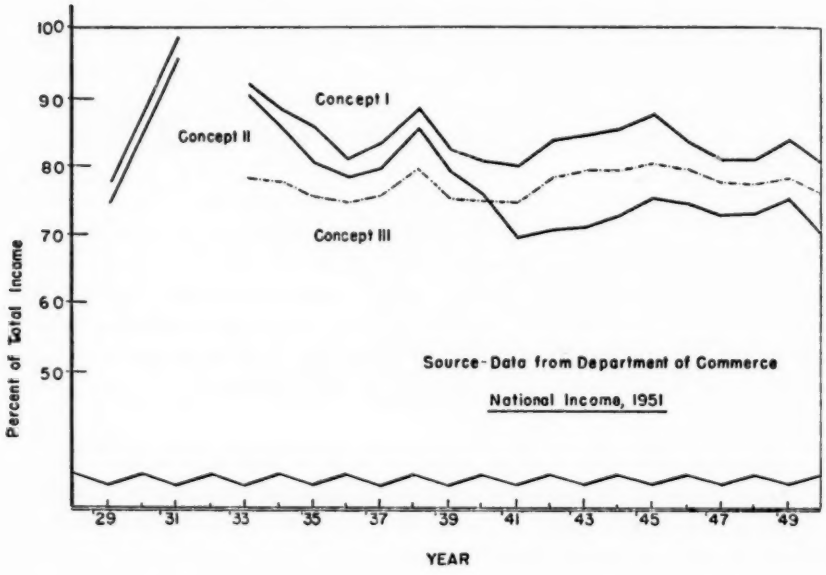
Department of Commerce data permit isolation and separate examination of labor and property shares in the corporate sector. The behavior of labor shares in accordance with varying concepts of income is set forth in Chart III; the data are shown in Appendix Table B. In general, Concepts I, II, and III are similar in definition to those utilized in Chart II.

Labor income, for all three concepts, is uniformly taken to include the compensation of employees of corporate enterprise. This compensation includes wages and salaries, employer contribution for social insurance and for private pension funds, and "other labor income." Property income in Concept I includes corporate profits after tax and net interest paid by corporations. For Concept II, property income also includes corporate profits taxes. For Concept III property income excludes corporate profits taxes, but includes corporate depreciation, for which data are available since 1933.[12]

[12] Corporate depreciation has been used because it is not possible to isolate capital consumption allowances for corporations in the Department of Commerce accounts.

As Chart III shows, fluctuations in labor shares in the corporate sector are considerably sharper than the fluctuations in labor shares in the economy as a whole. For example, the 1937–38 recession and the 1949 recession are considerably more pronounced. And, of course, the differences in behavior between Concept I and Concept II (corporate profits taxes included) are considerably sharper than for the economy

CHART III. Labor Shares in Total Income, Corporate Sector, the Primary Distribution of Income, 1929–1950.



as a whole. The inclusion of corporate depreciation (Concept III) seems to have an effect similar to the inclusion of consumption allowances for the economy as a whole.

The conclusions from Chart III are similar to those which may be derived from Chart II. Labor shares in the corporate sector were higher in the late Thirties than in 1929, and higher in World War II than in the late Thirties.

Postwar changes are more difficult to interpret. Where income is based on Concept I it would appear that the shrinkage in labor shares

It would be appropriate to include net rents in property income for the corporate sector-rents and royalties received by corporations less net rent paid on business property. Unfortunately, complete data on this type of income are not available for all years. However, in dollar amounts, and in its effect on relative shares, this omission is unimportant. For example, in 1942, net rents of corporations amounted to −$8 million. (U. S. Treasury Department, *Statistics of Income for 1942*, Part 2, pp. 328–29.)

was sufficiently pronounced to bring this sector back to about the same relative position that it occupied in the semi-prosperous years of the Thirties. However, where corporate depreciation is included (Concept III) the conclusion is comparable with that observed for the economy as a whole. The shrinkage in labor's relative position in the postwar years still left this sector in a slightly improved status as compared with the Thirties.

III

The measurements of labor and property shares in total income, as set forth in Part II, are based on the primary distribution of income, defined as payments or accruals from current economic activity. It is also useful to explore the behavior of relative shares when income is measured on the basis of claims. This requires attention to the appropriate classification of net capital gains and of transfer payments.

RELATIVE SHARES ON THE BASIS OF CLAIMS

Capital gains are rather obviously property income and have been classified as such here. Data on realized capital gains are taken from the Seltzer study.[13] Capital gains have been included in property income at net; that is, losses have been deducted from gains, and where, as in depression years, there are aggregate net losses, these are added algebraically to property income. Net gains realized by both individuals and corporations are included.

Transfer payments are classified as labor income. Again, it should be recalled that income is viewed here from the standpoint of the recipients. There can be little doubt that the recipients of labor income are, by and large, the recipients of transfer payments. In 1950, for example, $6,125 million of the $15,082 million of transfer payments represent benefits from social insurance funds. Military and veterans pensions and benefits, generally related to services previously performed, amounted to an additional $4,284 millions.

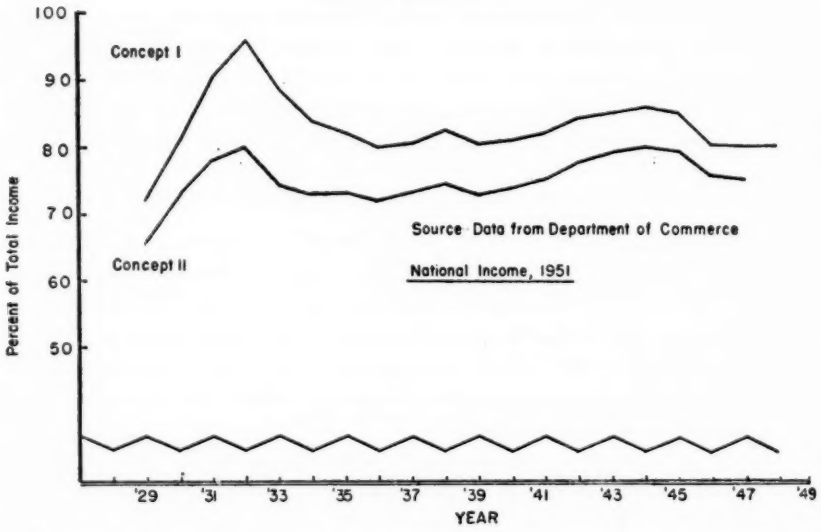
In combining transfer payments with labor income as defined in accordance with the primary distribution of income (Part II), it is necessary to deduct employer and employee contributions for social insurance and employer contributions for private pension funds from the total of labor income. This will avoid counting, as labor income, both employer and employee *contributions* to social insurance and pension funds and employee *benefits* from social insurance and pension funds.

[13] Lawrence H. Seltzer, *The Nature and Tax Treatment of Capital Gains and Losses* (National Bureau of Economic Research, 1951), pp. 367, 531.

This deduction has been made in Chart IV, and in Appendix Table C, which sets forth the data on which Chart IV is based.[14]

Chart IV shows labor's share in total income in accordance with the claims concept of income; transfer payments are included in labor income and net capital gains in property income. Concept III differs from Concept I in that capital consumption allowances are included in the property income segment.

CHART IV. Labor Shares in Total Income: the Claims Concept of Income, 1929–1950.



A comparison of Concept I, Chart II, with Concept I, Chart IV shows that the measurement of income on a claims basis produces results which are markedly different than the measurement of income on the basis of primary distribution. First, property income falls more sharply in the depression because of the impact of capital losses. Second, the wartime improvement in labor's relative position is still substantial, when compared with the Thirties, but is by no means as striking as when net capital gains and transfer payments are omitted. Third, labor's relative position in the postwar period appears to be almost identical with its relative position in the late Thirties. This is a reflection of the fact that transfer payments are less significant in rais-

[14] Employee benefits from social insurance funds are included in transfer payments, but benefits from private pension funds are not included, and cannot be isolated in the national income accounts. The procedure which has been adopted here will, therefore, tend to understate labor income to the extent of current benefits from private pension funds.

ing the labor share of total income in the postwar period share than are capital gains in raising the property share.

When capital consumption allowances are included (Concept III), it will be observed that the effect of the net gains and transfer payments is somewhat diluted. The behavior of Concept III in Chart IV is very similar to the behavior of Concept III in Chart II. The depression shrinkage in property incomes occasioned by capital losses is, of course, reflected in Chart IV, but the wartime improvement in labor's relative share appears to be less striking when net gains and transfer payments are included. For both the primary distribution of income and the claims concept of income, the inclusion of capital consumption allowances in property income tends to accentuate the improvement in labor's relative position over the whole period.

In conclusion, it would appear that the major differences between the measurement of labor and property shares in terms of the primary distribution of income and in terms of the claims concept of income are: 1) the claims concept of income shows labor's relative improvement to be more modest during World War II than does the primary distribution of income, and, 2) the claims concept shows labor's postwar relative position to be almost identical with its relative position in the late Thirties, while the primary distribution shows labor's relative position to be slightly improved in the postwar period.

LABOR AND PROPERTY SHARES IN THE TWENTIES

Data for the decade of the Twenties, comparable in refinement to the Department of Commerce accounts for 1929–50, are not available. However, the Kuznets estimates provide an adequate basis for broad comparison.

Chart V, for the years 1919–1929, is based on Kuznets data, with concepts generally similar to those described above in terms of the primary distribution of income.[15] As the chart indicates, from 1921 to 1929 there was remarkable stability in the labor-property division of total income.

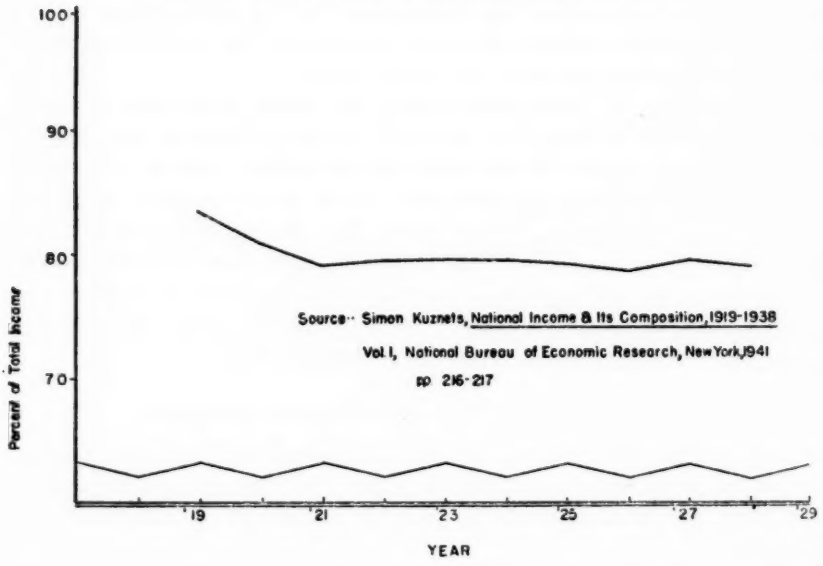
Data are not available for the estimation of labor and property shares in the Twenties in accordance with the claims concept of income; the necessary information on transfer payments is lacking. How-

[15] For a detailed discussion of these concepts see Simon Kuznets, *National Income and Its Composition, 1919–1938, Vol. I* (National Bureau of Economic Research, 1941), pp. 3–60, 215–65.

The concept of "total income" as used in Appendix Table D and Chart V differs from Kuznets' national income by the exclusion of government net savings. On the other hand, "total income" as used here differs from the primary distribution of income, as set forth in Chart I, principally in the exclusion of imputed rent on owned homes and food produced and consumed on farms.

ever, the Seltzer data on capital gains for individuals for these years would indicate that the stability exhibited in the labor-property relationship would be substantially modified by the addition of net gains to the property segment.[16] For example, in the years 1919–1923 net capital gains for individuals aggregate to an insignificant $30 million. But for the years 1924–1929 they aggregate to $15.8 billion. Total property income for these years amounted to $96.9 billion. For the single year 1928, in which net gains were most significant, their inclu-

CHART V. Labor Shares in Total Income, 1919–1929.



sion in property income would raise the relative share of property income in the total by about 20 per cent (four percentage points).

If it may be assumed that transfer payments during the years 1919–1929 were minimal, it is evident that a claims concept of income for the Twenties, in contrast with a primary distribution concept, would show a downward drift in labor shares beginning in the year 1924, and continuing through 1929.

IV

Before summarizing the general findings which emerge from this investigation it may be useful to explore briefly the relationship between

[16] See Seltzer, *op. cit.*, p. 367. Net gains of corporations are not available.

changes in labor and property shares of total income and resulting, or corresponding, changes in the distribution of personal income.

RELATIVE SHARES AND THE DISTRIBUTION OF INCOME

It is not possible to move directly from the measurement of distribution of income by type to the distribution of income by size class because many persons receive income from several sources. Nevertheless, the total income of upper income groups is composed more largely of property income than is the total income of lower income groups, and this fact establishes a rough linkage between changes in labor and property shares and changes in concentration in the distribution of income. On this basis it could be expected that an increase in property shares in total income would tend to increase concentration in the distribution of income. It could be expected that an increase in labor shares of total income would tend to reduce concentration in the distribution of income.

However, as Pechman has pointed out, a decrease in the ratio of wages to total income can be produced in two ways: first, since wages are a smaller proportion of income at higher income levels, the ratio of wages to total income might decrease as income recipients are shifted upward in the distribution; second, an increased concentration of income might reduce the share of total income going to wages.[17] Therefore, only when both total income and labor's share are concurrently increasing can it be certain that concentration is reduced; similarly, only when both total income and labor's share are falling can it be certain that concentration is increased.

It follows that, strictly speaking, only the World War II case shows an irrefutable instance in which the concentration of income is reduced. In all other instances in the years 1919–1950 where labor shares rise or fall, total income is moving in the opposite direction.

Nevertheless, the more generalized and less rigid hypothesis that an increase in labor shares is prima facie indication of reduced concentration seems to be generally borne out by the conclusions that have been reached on the basis of other methods of analysis. On the basis of his hypothesis it would be concluded that, for both the primary distribution of income and the claims concept of income, the distribution of income was less concentrated in the Thirties than in the Twenties and less concentrated during World War II than in the Thirties. Concentration increased in the 1946–1950 period as compared with World War

[17] Frank A. Hanna, Joseph A. Pechman, Sidney M. Lerner, *Analysis of Wisconsin Income* (National Bureau of Economic Research, 1948), pp. 69–75.

II, but the resulting degree of concentration was slightly less than in the late Thirties.

The Kuznets study of upper income group shares indicates that concentration decreased from 1939 to 1945 and increased somewhat after 1945.[18] The finding here that income distribution was less concentrated in World War II than during the middle Thirties is also in line with conclusions drawn from the surveys of family income for the years 1935–36, 1941, 1942 and 1945.[19]

On the other hand, the conclusion indicated in Chart II is that the 1946–1948 inflation tended to increase property shares and, therefore, in all probability, tended to increase concentration in the personal distribution of income. The survey of consumer finances conducted by the Board of Governors of the Federal Reserve System indicates that concentration increased from 1946 to 1947, but was reduced in the three succeeding years.[20] This is in conflict with the Chart II data, which shows concentration increasing continuously from 1946 through 1948, with a very modest reduction in concentration in 1949, and an increase again in 1950. There appears to be no way to reconcile these conflicting conclusions.

The finding in Chart II that labor shares tend to move against the business cycle, that is, increase during declines in economic activity and decrease as economic activity increases, would tend to indicate that the primary distribution of income is more concentrated during expansions than during contractions. The Kuznets data on concentration as revealed by changes in the shares of the upper one per cent of income recipients (where property income dominates) would appear to support this conclusion.[21] On the other hand, Mendershausen found that in 1933 income was somewhat more concentrated than in 1929.[22] However, these findings were based on examination of the top 30 per cent of income recipients. Mendershausen found that the top one per cent, where property income is wholly predominant, showed a reduction in concentration between 1929 and 1933.

It may be tentatively concluded that an increase in labor shares is

---

[18] Simon Kuznets, *Shares of Upper Income Groups in Income and Savings* (National Bureau of Economic Research, 1950), pp. 3–4.

[19] Morris A. Copeland, "The social and economic determinants of the distribution of income in the United States," *American Economic Review*, March 1947, 57. Also, for a comparison of 1929 and 1937 see Julius Wyler, "The share of capital in national income," *Social Research*, November 1943, 436–54.

[20] *Federal Reserve Bulletin*, August 1951, pp. 929.

[21] *Shares of Upper Income Groups in Income and Savings*, p. 33 (Panel A).

[22] Horst Mendershausen, *Changes in Income Distribution During the Great Depression* (National Bureau of Economic Research, 1946), pp. 68–80.

generally associated with an increase in equality in the distribution of income. However, the evidence does not suggest that this is an exact relationship which holds for moderate year to year changes in labor and property shares.

THE GENERALIZED BEHAVIOR OF LABOR AND PROPERTY SHARES

The relationships shown in Charts II, III and IV reveal a general pattern of behavior. As the level of economic activity increases, labor shares tend to decline. As the level of economic activity decreases, labor shares tend to increase. The only marked exception to this pattern is the case of World War II, when, as economic activity increased, labor shares also increased.

It would appear that there are a number of factors which contribute to this inverse relationship between labor shares and changes in economic activity.

As economic activity increases, property incomes can be expected to gain in relation to labor incomes by what may be called the *capacity effect*. As the output of firms increases beyond the break-even point, there is a concomitant increase in the volume of profits. The traditional spreading of overhead, with correspondingly lower unit labor costs, produces a larger return to property per unit of output as output increases. The growth in the relative shares of property income recipients is certainly not without limit; this limit is associated with a disproportionate increase in marginal costs.[23]

Where increases in economic activity are accompanied by inflationary conditions two additional factors will tend to increase the share of property income in relation to labor income. These may be termed the *lag effect* and the *compounding effect*. For present purposes the term inflation will be used to describe a condition in which demand for final product is increasing with the prices of goods and services generally rising, but with some prices rising more rapidly than others. That is, inflation is both an aggregative and a differential phenomenon.

The lag effect, which tends to increase property shares, is a corollary of the differential impact of price rises. A rising rate of effective demand confronts an output of goods and services which is increasing less rapidly than the increase in demand. The market for final product receives the full brunt of the higher demand; the increased demand for labor is secondary. The seller of commodities becomes the immediate beneficiary of the rising prices. Entrepreneurial shares increase and the

[23] For this point I am indebted to A. M. McIsaac.

increase raises the shares of property income in the total.[24] The lag effect is not independent of the capacity effect. Increases in price for final product will raise the break-even point and broaden the range of profitable operation.

Under inflationary conditions, a shift in relative shares of total income in favor of property income recipients will also result from what may be called the *compounding effect.* That which is compounded is savings, and the compounding occurs through its reinvestment.[25] Actually, this effect operates independently of price rises during any period in which operating losses and capital losses of firms are at a minimum. However, such a situation is characteristic of an inflationary period. Further, as entrepreneurial income expands, relatively, in an inflationary period, the tendencies toward compounding will be reinforced, if it may be assumed that the marginal propensity to consume out of entrepreneurial income is smaller than out of labor income. As long as inflationary conditions continue, with operating and capital losses at a minimum, those who have enjoyed the enhanced returns may further improve their economic position by earning income on income. The check on the compounding effect, is, of course, the appearance of operating and capital losses. A reduced rate of return on investment will limit its operation but not offset it.

The importance of each of these factors in the relative growth of property income under inflationary conditions will vary with the rate and duration of the inflation. Presumably, lag effect will be most pronounced when the rate of price increases is most rapid. Compounding effect, on the other hand, is likely to be of greatest importance when inflationary conditions have been sustained for long periods of time. The capacity effect will depend on the level from which increases in economic activity proceed and will be most important when increases in output are initiated from a point where many firms are operating close to break-even.

The operation of three factors—lag, compounding and capacity effect—will alter labor and property shares in accordance with changes in the level of economic activity and the level of prices. Lag and capacity effect, in particular, may be expected to be important in year to year fluctuations.

These three factors are independent of inter-industry shifts which

[24] The tendency of inflation to increase entrepreneurial shares has often been noted, but seldom explored in the literature. See Koopmans, *op. cit.;* and, of course, J. M. Keynes, *How to Pay for the War* (Harcourt, 1940), p. 6.

[25] Gordon Hayes, *Spending, Saving, and Employment* (Knopf, 1945), pp. 35–42.

may alter labor and property shares in total income, particularly over a period of years. Most writers in the field have stressed the importance of these shifts in altering labor shares. In examining the upward movement in relative labor income between 1919 and 1938 Kuznets found that relative increases in employee compensation were traceable almost wholly to production shifts from industries in which the ratio of wages and salaries to property income was low to industries in which the ratio of wages and salaries to property income was high. Kuznets did not find a measurable tendency for particular industries to vary their ratio of wages and salaries to property income over this whole period.[26] The Department of Commerce has also stressed the importance of interindustry shifts in the explanation of the larger employee shares in 1950 as compared with 1929.[27]

Unfortunately, it is not possible to disentangle statistically the year to year effects of the three factors noted above from the effects of interindustry shifts over a period of years on labor and property shares in total income.

Finally, it should be noted that the general conclusion which emerges from the examination of labor and property shares in accordance with the concepts developed here is that they are by no means stable, either from year to year or over a period of years. This conclusion differs rather sharply from that reached by analysts in the Department of Commerce who have stressed the stability in income shares, particularly in the employee compensation share of national income.[28] The explanation for these differing conclusions lies in the concepts which are employed. The Department of Commerce analysis of employee compensation is based on a concept of income which includes corporate profits taxes and inventory valuation adjustment. Both of these are excluded from Concepts I and III as developed here; the latter Concept also includes capital gains and transfer payments. The importance of the conceptual basis in the measurement of relative shares is well illustrated in the analysis of the corporate sector (Chart III and Appendix Table B). In 1929 the addition of corporate profits taxes to total income affects labor's share (Concept I) by only 2.4 percentage points. In 1950 this addition affects labor's share by 11.1 percentage points.

(*Continued on page 218*)

[26] Kuznets, *National Income and Its Composition*, pp. 241–50.

[27] Denison, "Distribution of National Income," pp. 16–23.

[28] *Ibid.*; Lawrence Groze, "Labor income in the postwar period," *Survey of Current Business*, May 1952, 7–13.

TABLE A

LABOR AND PROPERTY SHARES: THE PRIMARY DISTRIBUTION OF INCOME

1929–1950

(money figures in millions of dollars)

	1929	1930	1931	1932	1933	1934	1935	1936	1937
Property Sector:									
1. Rental Income	5,811	4,786	3,620	2,508	2,018	2,095	2,288	2,682	3,149
2. Corp. Profits After Tax	8,420	2,455	−1,283	−3,424	−362	977	2,259	4,273	4,685
3. Net Interest	6,541	6,176	5,938	5,430	5,010	4,750	4,539	4,474	4,376
4. Govern't Interest	983	964	1,084	1,141	1,170	1,230	1,141	1,101	1,204
5. Total Property (Concept I)	21,755	14,381	9,359	5,655	7,836	9,052	10,227	12,530	13,405
Labor Sector:									
6. Income of Unincorp.	13,927	10,963	8,214	4,921	5,207	6,603	9,858	9,942	12,246
7. Less: Inv. Val. Adj.	142	755	611	295	−525	−54	−50	−120	−29
8. Net Income of Unincorp.	13,785	10,208	7,603	4,626	5,732	6,657	9,908	10,062	12,275
9. Compensation of Employees	50,786	46,515	39,470	30,826	29,330	34,067	37,107	42,675	47,696
10. Total Labor (Concept I)	64,571	56,723	47,073	35,452	35,062	40,724	47,015	52,737	59,971
11. Total Property (Line 5)	21,755	14,381	9,359	5,655	7,836	9,052	10,227	12,530	13,405
12. Total Income (Concept I)	86,326	71,104	56,432	41,107	42,898	49,776	57,242	65,267	73,376
Ratios: Concept I									
13. Total Labor (Line 10)	74.8	79.8	83.4	86.2	81.7	81.8	82.1	80.8	81.7
14. Total Property (Line 11)	25.2	20.2	16.6	13.8	18.3	18.2	17.9	19.2	18.3
15. Total Property (Line 11)	21,755	14,381	9,359	5,655	7,836	9,052	10,227	12,530	13,405
16. Add: Corp. Profits Taxes	1,398	848	500	382	524	746	965	1,411	1,512
17. Total Property (Concept II)	23,153	15,229	9,859	6,037	8,360	9,798	11,192	13,941	14,917
18. Total Labor (Line 10)	64,571	56,723	47,073	35,452	35,062	40,724	47,015	52,737	59,971
19. Total Income (Concept II)	87,724	71,952	56,932	41,489	43,422	50,522	58,207	66,678	74,891
Ratios: Concept II									
20. Total Labor (Line 18)	73.6	78.8	82.7	85.4	80.7	80.6	80.8	79.0	80.1
21. Total Property (Line 17)	26.4	21.2	17.3	14.6	19.3	19.4	19.2	21.0	19.9
22. Total Property (Line 11)	21,755	14,381	9,359	5,655	7,836	9,052	10,227	12,530	13,405
23. Add: Cap. Cons. Allowances	8,816	8,747	8,312	7,663	7,245	7,218	7,369	7,684	7,972
24. Total Property (Concept III)	30,571	23,128	17,671	13,318	15,081	16,270	17,596	20,214	21,377
25. Total Labor (Line 10)	64,571	56,723	47,073	35,452	35,062	40,724	47,015	52,737	59,971
26. Total Income (Concept III)	95,142	79,851	64,744	48,770	50,143	56,994	64,611	72,951	81,351
Ratios: Concept III									
27. Total Labor (Line 25)	67.9	71.0	72.7	72.7	69.9	71.5	72.8	72.3	73.7
28. Total Property Line 24)	32.1	29.0	27.3	27.3	30.1	28.5	27.2	27.7	26.3

Source: Department of Commerce, *National Income, 1951.*

TABLE A (Continued)

1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	194	1949	1950
3,278	3,465	3,620	4,322	5,395	6,109	6,495	6,256	6,620	7,059	7,506	7,545	8,039
2,289	5,005	6,447	9,386	9,433	10,646	10,808	8,502	13,881	18,549	20,734	17,347	22,774
4,290	4,212	4,104	4,113	3,894	3,355	3,137	3,009	2,922	3,544	4,335	4,908	5,386
1,192	1,205	1,291	1,289	1,517	2,140	2,803	3,663	4,432	4,378	4,451	4,610	4,710
11,049	13,887	15,462	19,110	20,239	22,250	23,243	21,430	27,855	33,530	37,026	34,410	40,909
10,768	11,282	12,660	16,504	23,041	26,731	28,097	31,247	35,375	35,365	39,751	33,903	35,964
221	−166	−52	−644	−372	−154	−70	−113	−1,819	−1,547	−395	632	−1,552
10,547	11,448	12,712	17,148	23,413	26,885	29,067	31,360	37,194	36,912	40,146	33,271	37,516
44,747	47,820	51,786	64,280	84,895	109,212	121,163	123,026	117,098	127,988	140,166	139,887	153,333
55,294	59,268	64,498	81,428	108,308	136,097	150,230	154,386	154,292	164,900	180,312	173,158	190,849
11,049	13,887	15,462	19,110	20,239	22,250	23,243	21,430	27,855	33,530	37,026	34,410	40,909
66,343	73,155	79,960	100,538	128,547	158,347	173,473	175,816	182,147	198,430	217,338	207,568	231,758
83.3	81.0	80.7	81.0	84.3	85.9	86.6	87.8	84.7	83.1	83.0	83.4	82.3
16.7	19.0	19.3	19.0	15.7	14.1	13.4	12.2	15.3	16.9	17.0	16.6	17.7
11,049	13,887	15,462	19,110	20,239	22,250	23,243	21,430	27,855	33,530	37,026	34,410	40,909
1,040	1,462	2,878	7,846	11,665	14,406	13,525	11,215	9,583	11,940	13,028	10,989	18,593
12,089	15,349	18,340	26,956	31,904	36,656	36,768	32,645	37,438	45,470	50,054	45,399	59,502
55,294	59,268	64,498	81,428	108,308	136,097	150,230	154,386	154,292	164,900	180,312	173,158	190,849
67,383	74,617	82,838	108,384	140,212	172,753	186,998	187,031	191,730	210,370	230,366	218,557	250,351
82.1	79.4	77.9	75.1	77.2	78.8	80.3	82.5	80.5	78.4	78.3	79.2	76.2
17.9	20.6	22.1	24.9	22.8	21.2	19.7	17.5	19.5	21.6	21.7	20.8	23.8
11,049	13,887	15,462	19,110	20,239	22,250	23,243	21,430	27,855	33,530	37,026	34,410	40,909
7,992	8,101	8,440	9,294	9,981	10,680	11,887	12,410	12,163	14,845	17,612	19,058	21,177
19,041	21,988	23,902	28,404	30,220	32,930	35,130	33,840	40,018	48,375	54,638	53,468	62,086
55,294	59,268	64,498	81,428	108,308	136,097	150,230	154,386	154,292	164,900	180,312	173,158	190,849
74,335	81,256	88,400	109,832	138,528	169,027	185,360	188,226	194,310	213,275	234,950	226,626	252,935
74.4	72.9	73.0	74.1	78.2	80.5	81.0	82.0	79.4	77.3	76.7	76.4	75.5
25.6	27.1	27.0	25.9	21.8	19.5	19.0	18.0	20.6	22.7	23.3	23.6	24.5

TABLE B

LABOR AND PROPERTY SHARES IN THE CORPORATE SECTOR: THE PRIMARY DISTRIBUTION OF INCOME

1929–1950

(money figures in millions of dollars)

	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938
Property Sector:										
1. Corp. Profits After Tax	8,188	2,318	−1,279	−3,390	−360	917	2,100	4,169	4,563	3,042
2. Net Interest	1,617	1,742	1,802	1,759	1,715	1,707	1,626	1,557	1,463	1,482
3. Total Property (Concept I)	9,805	4,060	523	−1,631	1,355	2,624	3,726	5,726	6,026	4,524
4. Comp. of Employees	33,522	30,074	24,706	18,464	17,493	20,493	22,462	25,628	29,814	28,526
5. Total Property (Line 3)	9,805	4,060	523	−1,631	1,355	2,624	3,726	5,726	6,026	4,524
6. Total Income (Concept I)	43,327	34,134	25,229	——	18,848	23,117	26,188	31,354	35,840	33,050
Ratios: Concept I										
7. Total Labor (Line 4)	77.4	88.1	97.9	——	92.8	88.6	85.8	81.7	83.2	88.3
8. Total Property (Line 5)	22.6	11.9	2.1	——	7.2	11.4	14.2	18.3	16.8	11.7
9. Total Property (Line 5)	9,805	4,060	523	−1,631	1,355	2,624	3,726	5,726	6,026	4,524
10. Add: Corp. Profits Taxes	1,398	848	500	382	524	746	965	1,411	1,512	1,040
11. Total Property (Concept II)	11,203	4,908	1,023	−1,249	1,879	3,370	4,691	7,137	7,538	5,564
12. Total Labor (Line 4)	33,522	30,074	24,706	18,464	17,493	20,493	22,462	25,628	29,814	28,526
13. Total Income (Concept II)	44,725	34,982	25,729	——	19,372	23,863	27,153	32,765	37,352	34,090
Ratios: Concept II										
14. Total Labor (Line 12)	75.0	86.0	96.0	——	90.3	85.9	82.7	78.2	79.8	83.7
15. Total Property (Line 13)	25.0	14.0	4.0	——	9.7	14.1	17.3	21.8	20.2	16.3
16. Total Property (Line 5)					1,355	2,624	3,726	5,726	6,026	4,524
17. Add: Corp. Depreciation					3,498	3,366	3,357	3,287	3,338	3,352
18. Total Property (Concept III)					4,853	5,990	7,083	9,013	9,364	7,876
19. Total Labor (Line 4)					17,493	20,493	22,462	25,628	29,814	28,526
20. Total Income (Concept II)					22,346	26,483	29,545	34,641	39,178	36,402
Ratios: Concept III										
21. Total Labor (Line 19)					78.2	77.4	76.0	74.0	76.1	78.4
22. Total Property (Line 18)					21.8	22.6	24.0	26.0	23.9	21.6

Source: Department of Commerce, *National Income, 1951.*

TABLE B (Continued)

1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950
169	4,503	,042	4,821	6,213	9,155	9,208	10,408	10,515	8,274	13,427	17,825	19,868	16,506	21,683
557	1,493	,482	1,390	1,224	1,106	1,158	949	845	744	470	605	570	660	548
726	6,026	,524	6,211	7,437	10,261	10,366	11,357	11,360	9,018	13,897	18,430	20,438	17,166	22,231
628	29,814	,520	29,033	32,076	40,964	52,036	63,197	66,183	63,109	68,884	81,220	89,993	87,392	96,547
726	6,026	,524	6,211	7,437	10,261	10,366	11,357	11,360	9,018	13,897	18,430	20,438	17,166	22,231
354	35,840	,044	35,244	39,513	51,225	62,402	74,554	77,543	72,127	82,781	99,650	110,431	104,558	118,778
1.7	83.2	88.3	82.4	81.2	80.0	83.4	84.8	85.4	87.5	83.2	81.5	81.5	83.6	81.3
8.3	16.8	11.7	17.6	18.8	20.0	16.6	15.2	14.6	12.5	16.8	18.5	18.5	16.4	18.7
726	6,026	,524	6,211	7,437	10,261	10,366	11,357	11,360	9,018	13,897	18,430	20,438	17,166	22,231
411	1,512	,040	1,462	2,878	7,846	11,665	14,406	13,525	11,215	9,583	11,940	13,028	10,989	18,593
137	7,538	,564	7,673	10,315	18,107	22,031	25,763	24,885	20,233	23,480	30,370	33,466	28,155	40,824
528	29,814	520	29,033	32,076	40,964	52,036	63,197	66,183	63,109	68,884	81,220	89,993	87,392	96,547
765	37,352	,084	36,706	42,391	59,071	74,067	88,960	91,068	83,342	92,364	111,590	123,459	115,547	137,371
1.2	79.8	85.3	79.1	75.7	69.3	70.3	71.0	72.7	75.7	74.6	72.8	72.9	75.6	70.2
1.8	20.2	14.7	20.9	24.3	30.7	29.7	29.0	27.3	24.3	25.4	27.2	27.1	24.4	29.8
726	6,026	,524	6,211	7,437	10,261	10,366	11,357	11,360	9,018	13,897	18,430	20,438	17,166	22,231
387	3,338	,354	3,442	3,508	3,908	4,472	5,074	5,834	5,925	4,257	5,286	6,348	7,186	7,896
013	9,364	,878	9,653	10,945	14,169	14,838	16,431	17,194	14,943	18,154	23,716	26,786	24,352	30,127
28	29,814	,520	29,033	32,076	40,964	52,036	63,197	66,183	63,109	68,884	81,220	89,993	87,392	96,547
41	39,178	,398	38,686	43,021	55,133	66,874	79,628	83,377	78,052	87,038	104,936	116,779	111,744	126,674
.0	76.1	79.4	75.0	74.5	74.3	77.8	79.4	79.4	80.9	79.1	77.4	77.1	78.2	76.2
.0	23.9	20.6	25.0	25.5	25.7	22.2	20.6	20.6	19.1	20.9	22.6	22.9	21.8	23.8

TABLE C

LABOR AND PROPERTY SHARES: THE CLAIMS CONCEPT OF INCOME

1929–1950

(money figures in millions of dollars)

	1929	1930	1931	1932	1933	1934	1935	1936	1937
Property Sector:									
1. Total Property	21,755	14,381	9,359	5,655	7,836	9,052	10,227	12,530	13,4
2. Net Cap. Gains. Corps.	816[a]	−290	−1,404	−1,563	−1,423	− 55	231	439	1
3. Net Cap. Gains, Indiv.	2,893	−1,360	−2,718	−2,682	−1,403	−630	−89	589	−1
4. Total Property (Concept I)	25,464	12,731	5,237	1,410	5,010	8,367	10,369	13,558	13,4
Labor Sector:									
5. Total Labor Income	64,571	56,723	47,073	35,452	35,062	40,724	47,015	52,737	59,9
6. Less: Contrib. for Soc. Ins.[b]	371	377	383	391	388	427	462	767	1,5
7. Total	64,200	56,346	46,690	35,061	34,674	40,297	46,553	51,970	58,4
8. Transfer Payments	1,499	1,544	2,673	2,152	2,113	2,193	2,389	3,520	2,4
9. Total Labor (Concept I)	65,699	57,890	49,363	37,213	36,787	42,490	48,942	55,490	60,8
10. Total Property (Line 4)	25,464	12,731	5,237	1,410	5,010	8,367	10,369	13,558	13,4
11. Total Income (Concept I)	91,163	70,621	54,600	38,623	41,797	50,857	59,311	69,048	73,8
Ratios: Concept I									
12. Total Labor (Line 9)	72.1	82.0	90.4	96.3	88.0	83.5	82.5	80.4	81.9
13. Total Property (Line 10)	27.9	18.0	9.6	3.7	12.0	16.5	17.5	19.6	18.1
14. Total Property (Line 4)	25,464	12,731	5,237	1,410	5,010	8,367	10,369	13,558	13,4
15. Add: Cap. Cons. Allowances	8,816	8,747	8,312	7,663	7,245	7,218	7,369	7,684	7,9
16. Total Property (Concept III)	34,280	21,478	13,549	9,073	12,255	15,585	17,738	21,242	21,4
17. Total Labor (Line 9)	65,699	57,890	49,363	37,213	36,787	42,490	48,942	55,490	60,8
18. Total Income (Concept III)	99,979	79,368	62,912	46,286	49,042	58,075	66,680	76,732	81,8
Ratios: Concept III									
19. Total Labor (Line 17)	65.7	72.9	78.5	80.4	75.0	73.2	73.4	72.3	73.6
20. Total Property (Line 18)	34.3	27.1	21.5	19.6	25.0	26.8	26.6	27.7	26.4

Source: Department of Commerce, *National Income, 1951:* capital gains from Lawrence H. Seltzer, *The Nature and Tax Treatment of Capital Gains and Losses* (National Bureau of Economic Research, 1951), pp. 367, 531.

[a] Estimated from Department of Commerce, *National Income, 1951*, Table 38, p. 202.

[b] Includes employer contributions for social insurance and for private pension funds and employee contributions for social insurance.

TABLE C (Continued)

1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950
11,049	13,887	15,462	19,110	20,239	22,250	23,243	21,430	27,855	33,530	37,026	34,410	40,909
75	76	−672	−956	−176	−160	63	641	1,269	924	975[a]		
−341	−269	−437	−896	−380	1,057	1,602	4,267	6,644	4,450[a]	4,325[a]		
10,783	13,694	14,353	17,258	19,683	23,147	24,908	26,338	35,768	38,904	42,326		
55,294	59,268	64,498	81,428	108,308	136,097	150,230	154,386	154,292	164,900	180,312	173,158	190,849
2,122	2,292	2,452	2,967	3,715	4,908	5,897	7,019	7,222	7,268	7,084	7,698	9,379
53,172	56,976	62,046	78,461	104,593	131,189	144,333	147,367	147,070	157,632	173,228	165,460	181,470
2,834	2,963	3,119	3,119	3,150	2,971	3,597	6,179	11,420	11,803	11,285	12,352	15,082
56,006	59,939	65,165	81,580	108,043	134,160	147,920	153,546	158,490	169,435	184,513	177,812	196,553
10,783	13,694	14,353	17,258	19,683	23,147	24,908	26,338	35,768	38,904	42,326		
66,789	73,633	79,518	98,838	127,726	157,307	172,828	179,884	194,258	208,339	226,839		
83.9	81.4	81.9	82.5	84.6	85.3	86.1	85.4	81.6	81.3	81.3		
16.1	18.6	18.1	17.5	15.4	14.7	13.9	14.6	18.4	18.7	18.7		
10,783	13,694	14,353	17,258	19,683	23,147	24,908	26,338	35,768	38,904			
7,992	8,101	8,440	9,294	9,981	10,680	11,887	12,410	12,163	14,845			
18,775	21,795	22,793	26,552	29,664	33,827	36,795	38,748	47,931	53,749			
56,006	59,939	65,165	81,580	108,043	134,160	147,920	153,546	158,490	169,435			
74,781	81,734	87,958	108,132	137,707	167,987	184,715	192,294	206,421	223,184			
74.9	73.3	74.1	75.4	78.5	79.9	80.1	79.8	76.8	75.9			
25.1	26.7	25.9	24.6	21.5	20.1	19.9	20.2	23.2	24.1			

[a] Estimated from Treasury Department, *Statistics of Income*. It has been assumed, for 1947 and 1948, that short term gains are negligible. Therefore, the net gains as reported for tax purposes have been doubled to produce the estimates here.

TABLE D

LABOR AND PROPERTY SHARES, 1919–1929

(money figures in millions of dollars)

	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929
Property Sector:											
1. Rent	4.0	4.3	4.5	4.9	5.2	5.6	5.5	5.1	5.1	4.9	4.9
2. Dividends	2.9	3.2	3.0	3.0	3.8	3.8	4.4	4.7	5.1	5.5	6.3
3. Interest	3.2	3.7	3.9	4.0	4.2	4.4	4.6	4.7	4.9	5.3	5.6
4. Corp. Net Savings	1.0	2.2	.71	.23	.97	.42	.83	2.3	.56	.92	1.5
5. Total Property	11.1	13.4	12.11	12.13	14.17	14.22	15.33	16.8	15.66	16.62	18.3
Labor Sector:											
6. Wages & Salaries	36.7	43.3	34.9	36.4	42.7	42.7	44.4	47.4	47.8	48.7	51.5
7. Other Pay. to Empl.	.43	.57	.60	.60	.62	.62	.61	.62	.65	.66	.69
8. Entre. Withdr.	11.8	13.5	10.3	10.8	11.3	11.9	12.5	12.5	12.6	12.9	13.4
9. Entre. Savings	5.5	1.6	.63	−.09	1.2	.87	1.6	2.1	1.1	.91	1.1
10. Total Labor	54.43	58.97	46.43	47.71	55.82	56.09	59.11	62.62	62.15	63.17	66.69
11. Total Property (Line 5)	11.1	13.4	12.11	12.13	14.17	14.22	15.33	16.8	15.66	16.62	18.3
12. Total Income	65.53	72.37	58.54	59.84	69.99	70.31	74.44	79.42	77.81	79.79	84.99
Ratios:											
13. Total Labor (Line 10)	83.1	81.5	79.3	79.7	79.8	79.8	79.4	78.8	79.9	79.2	78.5
14. Total Property (Line 11)	16.9	18.5	20.7	20.3	20.2	20.2	20.6	21.2	20.1	20.8	21.5

Source: Simon Kuznets, *National Income And Its Composition, 1919–38*, Vol. I (National Bureau of Economic Research, 1941), pp. 216–17.

SUMMARY

The major conclusions which may be drawn from the foregoing investigation of labor and property shares in total income are the following:

1. Since the Twenties labor shares in total income have tended generally to increase. The recovery period of the Thirties shows labor shares higher than in the Twenties, and higher in World War II than in the Thirties. However, the postwar period has tended to reduce labor shares very nearly to their position in the late Thirties.

2. The available data on the distribution of income by size class, when compared with changes in labor and property income shares, would seem to indicate that increased labor shares are generally associated with reduced concentration in the distribution of income.

3. As economic activity declines, labor shares tend to increase. As economic activity increases, labor shares tend to decline. A decline in labor shares (except for World War II) seems to be accentuated by inflationary conditions.

4. The measurement of income on the basis of claims, rather than on the basis of primary distribution, alters somewhat the relative change

in labor and property income since the Twenties, but does not alter the relationship between changes in economic activity and changes in labor's relative share.

5. Where the measurement of income includes capital consumption allowances, relative shares appear to be considerably more stable between prosperity and depression.

The conclusion which is of most general theoretical interest is the third—that increases in economic activity are associated with a reduction in labor shares. This point is particularly significant if a reduction in labor shares is linked to increased concentration in the distribution of income. For this will mean that, if the marginal propensity to consume of property income recipients is lower than the marginal propensity to consume of labor income recipients, increases in economic activity will tend to increase savings ratios and, therefore, increase the volume of investment necessary to maintain the higher level of income. Since inflation appears to accentuate the shift to property income, it may be said that the inflationary process itself makes it more difficult to maintain prevailing levels of national income. Unfortunately, our present knowledge of consumption ratios by income class does not permit a final assessment of the significance of this point. We do not now know the range of variation in the marginal propensity to consume by income class or by type of income.

It is also of considerable interest that the "normal" pattern of relationship between labor and property shares and changes in economic activity was substantially altered during World War II. The increase in economic activity during this period was associated with an increase, not a decrease, in labor's share of total income. A comparison of this experience with the experience of the years 1946–48 and 1950, when labor shares were declining, suggests that the effective stabilization program of the war years may be responsible for the large increases in labor shares of total income in these years.

THE POST-ENUMERATION SURVEY OF THE 1950 CENSUS: A CASE HISTORY IN SURVEY DESIGN

ELI S. MARKS, W. PARKER MAULDIN AND HAROLD NISSELSON
Bureau of the Census

THE primary purpose of the present paper is to trace the development of the Post-Enumeration Survey, which was designed to measure error in the 1950 Censuses, and to outline the reasons for the decisions made and the alternatives considered in designing this survey. This paper points up the extent to which decisions on survey design have to be reached on the basis of intuition and opinion, due to the absence of any satisfactory objective data. Such analysis directs attention to the major gaps in the knowledge of survey technique—the spots where there is need to supersede art with science—and may raise as subjects for investigation, points which have been implicitly accepted as axiomatic. For example, discussions of field methods usually devote considerable space to improving the selection and training of interviewers but practically no attention is given to the problem of whether the gains from such improvements are commensurate with the effort and expenditure involved.

The existence of errors in a Census is readily apparent in view of the vastness of the undertaking, with its necessarily large and (in large part) inexperienced enumerating staffs who are called on to obtain facts on a wide variety of subjects from a group of respondents who may be given only a partial understanding of what is wanted and who may have a very incomplete knowledge of the information requested. In addition to these and other problems in the collection of the data, there are also human and mechanical failures in the processing of the data that are not always detected.

The recognition of the existence of such errors is not new. For many years, the Census volumes have contained introductory statements pointing out inconsistencies and possible sources of error in the statistics. Also, analyses of error in, and questions regarding the accuracy of, Census data can be found in numerous articles, a few of which are listed at the end of this paper.

What is relatively new is the attempt to provide a measure of the errors on the basis of an independent enumerative check. Checks of this sort were conducted in connection with the 1945 Census of Agriculture, the 1947 Census of Manufactures, and the 1948 Census of Business. These experiences formed the basis for planning the 1950 Census Post-Enumeration Survey.

The design of the Post-Enumeration Survey was premised on a need for a high level of accuracy—a level substantially higher than required or attained in the Census. Since the Census was taken with considerable attention to securing accurate results, such a requirement was quite rigorous. On the other hand, it was agreed that perfection was neither attainable, nor necessary, in order for the results to be useful.

Not only was there the desire to measure with precision the amount of error, but also, to the extent possible, to determine the reasons for error. This second feature is particularly significant since exploration of the sources of error should provide our most valuable guide to improving future surveys. It can be seen that a study of "response variation" would not answer these purposes. Measures of response variation—i.e., the differences which occur when identical questions are asked under (presumably) identical conditions—help to point out weaknesses in our data. They do not, however, throw much light on how to correct these weaknesses. The fact that two responses differ indicates that one (or both) of the responses is in error but does not indicate which one (if either) is correct.

A brief statement indicating the organization of the Bureau may help in understanding how the Post-Enumeration Survey was planned and conducted. Under the Director and Deputy and Assistant Directors, there are a number of Divisions. Those concerned in the Decennial Census, and hence in the Post-Enumeration Survey, were the Agriculture Division, the Field Division, the Geography Division, and the Population and Housing Division. The Office of the Assistant Director for Statistical Standards was given the responsibility for coordinating these interests. All major issues were reviewed and decisions formulated or revised in a committee consisting of representatives of all these organizational units. In addition, valuable advice and assistance was received from others outside the Bureau. This paper is, necessarily, a report not of work done by a single individual, but of decisions and plans formulated by a group.

In preparing the paper, minutes of meetings and other working records have been consulted but many of the gaps in these records (particularly in the area of alternatives considered) had to be filled in from memory. Of necessity, the description of the survey is telescoped and many important details have been omitted. A complete description of the checking methods used in the Census, or even of the Post-Enumeration Survey itself has not been attempted. Materials of this type will be incorporated in the publications giving the results of the Census and Post-Enumeration Survey.

SOME MAJOR FACTORS IN THE SURVEY DESIGN

The Post-Enumeration Survey was primarily concerned with:

A. Coverage errors, consisting of:
   (1) Omissions of persons, households, dwelling units, and farms.
   (2) Duplications of persons, households, dwelling units, and farms.
   (3) Inclusion of persons, households, dwelling units, or farms, but in the returns for the wrong area.

B. Content errors, consisting of omissions or incorrect entries as responses to the specific questions on the schedules.

One of the first problems that had to be met was that of sample design. In designing the sample several factors had to be considered:

(1) *Measurement of omissions.* For most purposes, the Census provides a satisfactory listing of the population. It cannot, of course, provide a satisfactory listing for the purpose of measuring its own completeness and, in general, any other source would be even less satisfactory. Thus, it was necessary to find some technique for drawing a sample of those units (persons, households, families, farms) which were not listed in the Census. The only technique available appeared to be to draw a sample of small areas (segments) and make special efforts in the Post-Enumeration Survey to list all units which were in the segment at the time of the Census. It could then be determined which units were enumerated by comparing the Post-Enumeration Survey segment listings with the Census listings.

(2) *Measurement of duplications and inclusions in the wrong area.* Checking "coverage" of the Census involves both the completeness of the enumeration and its accuracy. The segment sample would provide a measure of completeness. For checking accuracy of coverage, however, it was also necessary to determine how many units had been erroneously included in the Census listing. To identify erroneously-included units by comparison of the Post-Enumeration Survey segment listings with the Census listings would have required constructing a Census listing for the segment. While all units in the Census can be identified as being located in a particular Enumeration District (the area assigned to a particular enumerator), the Post-Enumeration Survey segments were considerably smaller than an Enumeration District. In many cases (approximately 20 per cent of the total), addresses shown within Enumeration Districts are not specific enough to identify the location by segment. Thus, any segment listing constructed from the Census would inevitably contain units which had been enumerated

in the correct Enumeration District but which were not actually located in the segment to which they had been assigned. This would result in an overestimate of the number of erroneously included units and would also increase the sampling variance of estimates of net coverage error (i.e., the difference between erroneously omitted and erroneously included units).

An alternative would have been to estimate the number of erroneously included units by subtracting an estimate of the number of units in the Census which were enumerated in the correct place (as determined from the Post-Enumeration Survey interviewers' segment listings) from the total number of units enumerated in the Census. Assuming a fixed total expenditure, however, an estimate obtained in this manner would be subject to a considerably larger sampling error than one obtained by sampling the Census listings and directly identifying the units erroneously included.

In addition, the indirect technique would not give any clues regarding the *reasons* for over-enumeration, i.e., how many over-enumerations were a result of carelessness, how many were a result of complexities or ambiguities in the rules determining who should be enumerated, etc. The segment technique would require a third visit to determine the reasons for over-enumerations.

Thus, it was decided to use two samples, a segment sample for units which were missed and a sample of Census listings to measure erroneous enumerations. Since an interview to check on over- or under-enumeration was necessary for every unit in both samples, it would have been desirable to overlap the samples to the maximum extent possible. Complete overlap would have required that one be able to identify in advance the segment in which each Census listing was located. Although this was impossible for the reasons discussed above, an effort was made to obtain a high degree of overlap.

(3) *Measurement of content error.* With this design, the units which were defined correctly (e.g., a farm which existed, which should have been enumerated and which was enumerated in the correct area, but not necessarily enumerated with the proper characteristics) might have been obtained from either or both samples. Information from such units was desired to determine whether their characteristics had been correctly reported.

Where overlap of the two samples was not achieved, it was necessary either to take complete interviews on all units in both samples and adopt a complex weighting system or to restrict the check on accuracy of characteristics for correctly included units either to those in the "seg-

ment sample" or those in the "list sample." Complete interviews for all units would, of course, have been more expensive than restricting complete interviews to only one of the samples (particularly when the cost of weighting is included in the expenditures). The additional expenditure would not have been justified by the increase in information obtained. Because it was considered desirable to furnish the check interviewer with the information obtained in the Census for a given unit (see the discussion below of "Transcription"), it was decided to restrict complete interviews to units in the sample of Census listings since they could be identified in advance. It was also necessary, of course, to obtain complete information from the "missed" units in the segment sample.

(4) *Length of interview.* In designing the Post-Enumeration Survey interview, emphasis was placed on investigating each variable thoroughly even if this meant restricting the number of variables to be checked. After the elimination of a great many Census variables from the Post-Enumeration Survey, it was found that the interview would still be quite extensive if the remaining variables were to be checked thoroughly. This did not present any serious problem except in farm households where it might be necessary to check the definition of the dwelling unit and of the farm, the completeness and correctness of the listing of persons, the housing characteristics of the dwelling unit, the farm characteristics, and the characteristics of each member of the household. Experience in a pretest indicated that a full interview for farm households would average $1\frac{3}{4}$ hours and that the upper quartile would exceed two hours. It was felt that interviews of this length would be undesirable.

The alternatives were (a) to drop additional items from the check, (b) to check some (or all) of the items less thoroughly, or (c) to check some of the items in one sample and others in another—in particular to separate the farm check from the check on population and housing characteristics. Dropping items from those originally selected was not an acceptable solution since appreciable savings in interview time could be made only by eliminating items which were of primary importance. A less thorough check was also undesirable—more, rather than less, investigation might be necessary to insure results of the required accuracy. Checking different items in different samples was feasible but might mean more complex instructions to the field staff, complications in processing the data, increased cost or increased sampling error and, possibly, increase in the error from sources other than sampling varia-

tion. Despite its disadvantages, this alternative did permit cutting the length of interviews in farm households without affecting the interviews for nonfarm households.

An additional factor in the decision was the complexity of the job facing the interviewer—he would have to master the definition of a dwelling unit and of a place requiring a farm check; the definitions of more than 100 different Census variables; the problems of defining "usual residence" and "farm headquarters"; the special procedure for occupied living quarters which did not meet the dwelling unit definition and for other places which were by definition not dwelling units; the use of maps and aerial photographs; the preparation of rough detail maps; the methods of recording on seven different forms; how to interview and how to probe for causes of discrepancies between his own results and those obtained by the Census enumerator; and the various special modifications of the procedures required by subsidiary studies which were linked to the Post-Enumeration Survey.

In the light of these considerations, it was decided to draw three samples in rural areas: the segment sample to be checked to locate missed households and farms; a sample of Agriculture Census listings to be checked for over-enumeration and for farm characteristics (the farm sample was overlapped with the segment sample as much as possible); and an independent sample of listings from the Population and Housing Census to be checked for over-enumeration on population and housing characteristics. The segment and farm samples were assigned to rural interviewers and the household sample to urban interviewers. The split was not made in urban areas where farm checks would be relatively infrequent.

Although the decision was predicated in part on the assumption that length of time required for a complete interview in farm households would be excessive, this is by no means an established fact. The length might have been feasible and, even if it were not, there were other means of reducing interview length—actually it was later decided (see below) to subsample individuals within households which cut the length of interview per household, but in view of the other considerations and the pressure of circumstances, the decision to check agriculture items and population and housing items in separate samples was not reexamined.

(5) *Selection, training and supervision of interviewers.* The objectives of the Post-Enumeration Survey required that interviewers be carefully selected, well-trained and competently supervised. With a fixed

budget, increasing the number of interviewers meant either reducing the training and supervision expenditures per interviewer or cutting at some other point.

INITIAL DECISIONS ON SURVEY DESIGN

As a first step in design of the sample, comparisons were made of estimated costs of different designs. Preliminary computations pointed to use of a design involving 300 primary sampling units. Drawing of a sample on this basis was, therefore, initiated. A two-stage sample was drawn, drawing primary sampling units consisting of a county or a group of contiguous counties and within the primary sampling unit selecting segments containing 6 households in urban areas or 10 households in rural areas. Later recalculation as shown in Table I indicates minimum cost for 200 primary sampling units. With the rough assumptions made about costs and intraclass correlations, the differences between 200 and 300 primary sampling units (with the same staffing pattern) are not of magnitude sufficient to form a basis for choice and the design with 300 primary sampling units actually used is presumed to be near the minimum cost design.

Final computations made to compare different survey designs are summarized in Table I. Designs 1 to 7 assume that all interviewers would conduct population, housing and agriculture interviews. Designs 8 to 13 assume that one set of interviewers (urban) would conduct population and housing interviews and another set (rural) would conduct agriculture interviews, within the same sample of primary sampling units but in different segments. Designs 2, 3, 6, and 7 show differences in cost associated with different numbers of primary sampling units in the situation where each interviewer would be assigned one primary sampling unit. Designs 1 and 4 indicate the case where each primary sampling unit is split between interviewers. In this case, there is a substantial increase in training and supervision costs with no compensating economies (under the assumptions made). In Design 5 there is only 1 interviewer for each 2 primary sampling units, which involves, therefore, an increase in travel and enumeration costs (mainly due to the fact that the interviewer would be paid subsistence allowance while working away from his home primary sampling unit) but this increase is compensated by a decrease in the training and supervision costs. Designs 8 through 13 show similar variations in the more complex situation where two types of interviewers are used and the ratios of rural interviewers per primary sampling unit and urban interviewers per primary sampling unit are varied.

TABLE I

ESTIMATED COSTS OF ALTERNATIVE POST-ENUMERATION SURVEY DESIGNS HAVING THE SAME VARIANCE

Design	Number of sample primary sampling units (1)	Number of sample households (2)	Number of interviewers (3)	Training cost including travel to and from training (4)	Within segment cost—including travel & enumeration cost (5)	Cost of travel between segments within primary sampling units (6)	Total cost of daily travel to and from beginning point of enumeration (7)	Travel cost between primary sampling units (9)	Processing costs for sample designation, transcription, editing, coding and punching (9)	Cost of supervision (10)	Total cost (11)
1	100	29,056	200	30,238	47,071	9,767	27,802	—	92,011	38,358	245,247
2	200	25,904	200	30,688	41,965	13,042	25,777	—	82,029	36,916	230,417
3	300	25,000	300	42,801	40,500	15,692	25,588	—	79,167	44,070	247,818
4	300	25,000	600	75,831	40,500	15,692	25,588	—	79,167	67,244	304,022
5	300	25,000	150	24,547	58,000	19,296	31,460	3,210	79,167	33,437	249,117
6	400	24,572	400	54,500	39,807	17,964	25,730	—	77,811	51,777	267,589
7	600	24,158	600	77,181	39,136	21,815	26,247	—	76,500	67,649	308,528
			Urban Rural								
8	200	25,904	200 100	42,464	48,829	18,833	27,862	2,621	82,029	43,997	266,635
9	300	25,000	200 100	42,801	57,917	25,343	31,525	6,420	79,167	44,728	287,901
10	300	25,000	240 60	42,801	55,350	24,914	30,627	6,420	79,167	44,728	284,007
11	300	25,000	150 150	42,801	61,125	25,880	32,648	6,420	79,167	44,728	292,769
12	300	25,000	400 200	75,831	36,333	22,117	27,482	2,140	79,167	67,847	310,917
13	600	24,158	400 200	77,181	55,966	35,232	32,911	9,079	76,500	68,549	355,418

1. For designs 1 through 6, it is assumed that no distinction is made between urban and rural interviewers.
2. The variances were estimated assuming primary sampling units drawn with probability proportional to their 1940 populations and sampling of segments (of six households each) within primary sampling units at rate inversely proportional to the 1940 population of the primary sampling unit. Variances were assumed to be independent of the number of Post-Enumeration Survey interviewers.
3. To equate the variances of the different designs, the intraclass correlation between segments within primary sampling units was taken as .005 and the level of accuracy obtained with 300 primary sampling units and 25,000 households was adopted as a standard.
4. In estimating travel distances, it was assumed that segments would have a random distribution within primary sampling units, that primary sampling units would have a random distribution over the whole United States and that the interviewers' homes would be uniformly distributed among the sample primary sampling units. It was also assumed that there would be a training center for every 20 interviewers, the centers being uniformly distributed over the whole United States. Distances were taken proportional to the square root of the area involved (the whole United States, the primary sampling unit or the training center area) and to the square root of the number of random points which had to be covered during travel.
5. An average of approximately 1½ hours per household was assumed for time of enumeration and within segment travel. Where interviewers were specialized, 25 per cent was added to this time for the rural enumeration.
6. It was assumed that mileage would be paid for at 7 cents per mile, that travel to the training center and between primary sampling units would take place on weekends or during other "nonworking time."
7. Interviewer's pay was taken at $1.30 per working hour and per diem (maintenance payments while in travel status) at $9.00 per day.
8. Cost of supervision was based on an average of 1 supervisor for 10 interviewers, the supervisors to work 5 weeks longer than the interviewers. During the interviewing it was assumed that supervisors would spend half their time in travel status. Supervisor's pay was assumed to be $2.25 per working hour.

Within the selected primary sampling units, it was decided to use segments containing about 6 households in urban areas and about 10 households in rural areas. From a standpoint of sampling variance, this size of segment may be too large for maximum efficiency. On the other hand, smaller segments would have meant problems in obtaining clear definition of boundaries, thus introducing "coverage" errors into the Post-Enumeration Survey.

SPECIALIZATION OF INTERVIEWERS

The decision to separate the agriculture check from the check of population and housing and to specialize the interviewers has been noted above. While estimates of the cost of such specialization are available (see Table I), no estimates of the gains have been made—Table I presents cost for a fixed *sampling* error but does not take into account other types of error. It was anticipated that the complexity of the interviewer's job might result in some confusion on his part so that his performance would be seriously impaired. It was also felt that, within the available training budget, intensive instruction on all phases was impossible unless the scope of the job was reduced. Thus, the anticipated gain from specializing interviewers was in accuracy and in reduction of the cost of correcting inadequate work. The magnitude of these gains could not be estimated even roughly in advance, and the survey itself yields only impressions on this point. While the decision made assumes implicitly that the gain was worth an estimated 20 to 30 thousand dollars additional cost (the difference in cost as shown in Table I is approximately $35,000 to $45,000, but part of this is chargeable to specialization and part to the additional sample in rural areas), this is a purely intuitive assumption. Subjective impressions indicate that interviewer training and initial performance were better than might have been obtained without specialization.

TRANSCRIPTION

Obviously the sample design affected (and was affected by) almost all phases of the Post-Enumeration Survey. Another problem which had considerable ramifications was the question of whether or not the re-interviewer should or should not be provided with the results of the original enumeration.

There are definite advantages (at least in theory) to giving the checkers the original data: (1) it permits the re-interviewer to check his own answers against the original entry and to determine more positively what the correct answer is; and (2) important leads can be obtained for

the improvement of future surveys by having the re-interviewer set down as much information as possible about the reasons for any errors found. On the other hand, there may be danger of minimizing the extent of error due to the interviewer "confirming" the original entry even where it is wrong. In an early Census pretest there was evidence that this occurred. However, the post-enumeration survey interviewers in

TABLE II

PER CENT OF POPULATION AND HOUSING SCHEDULES SHOWING A DIFFERENCE BETWEEN ORIGINAL AND RE-INTERVIEW ENTRIES, POST-ENUMERATION SURVEY OF MAY 1949 CENSUS PRETEST

Item	Group A* (Census Entries Available to Post-Enumeration Survey Interviewers)	Group B* (Census Entries Not Available to Post-Enumeration Survey Interviewers)
Age	29	25
Place of Birth	†	†
Work Status	11	8
Occupation	29	31
Industry	14	19
Class of Worker	25	23
Highest School Grade	41	38
Condition of Unit	19	18
Water Supply	6	15
Bathtub or Shower	2	8
Tenure	11	18

* Percentages for housing items based on 54 Group A and 60 Group B cases, for highest school grade on 37 Group A and 34 Group B cases, and for other items on 100 to 200 cases. Figures for Group A represent the differences remaining after comparison and reconciliation between the re-interviewer's initial entry and the figure transcribed from the original enumeration. Blank or omitted entries were not counted as differences.

† Less than one per cent.

this pretest were not specially selected or trained, and little restriction was placed on their use of the original data.

The technique finally evolved for use of the original data in the Post-Enumeration Survey gave the interviewer the information obtained in the Census for the "characteristics" of persons, households, and farms. He was not to look at this information until he completed the independent interview and, to reinforce this instruction, the original Census entries were transcribed to a page which folded in so that inadvertent examination of the entries was practically impossible. After the inter-

viewer completed his independent questioning, he compared his entries with the Census data and attempted to reconcile the differences by further questioning of the respondent.

In a test of the effectiveness of this procedure, conducted following a

TABLE III

NUMBER OF AGRICULTURE QUESTIONNAIRES SHOWING A DIFFERENCE BETWEEN ORIGINAL AND RE-INTERVIEW ENTRIES, POST-ENUMERATION SURVEY OF MAY 1949 CENSUS PRETEST

Item	Group A (Census entries available to Post-Enumeration Survey interviewers)			Group B (Census entries not available to Post-Enumeration Survey interviewers)	
	No. of reports	No. of differences		No. of reports	No. of differences
		Before reconciliation*	After reconciliation*		
Corn					
acres	24	14	12	22	15
bushels	23	19	18	22	17
Cotton					
acres	23	12	10	15	9
bales	23	10	9	15	8
Fruit and nut trees	31	7	7	24	5
Number of cattle	27	9	5	22	6
Chicken eggs, sold					
dozens	18	16	11	16	14
value	18	16	11	16	15
Work off the farm	33	14	14	26	11

* Figures in the column "Before reconciliation" represent differences before the re-interviewer had examined the data from the original enumeration. Figures in the column "After reconciliation" indicate number of differences remaining after comparison and reconciliation between the re-interviewer's initial entry and the figure transcribed from the original enumeration.

Census pretest in May 1949, the interviewer was not given the original Census information for one-half of the sample cases. The results (see Tables II and III) indicated that the interviewer's initial entries showed about the same proportion of differences from the Census entries regardless of whether or not he was given the original Census information. In some cases, however, the reconciliation disclosed errors in the inter-

viewer's initial entry which he could then correct. Because of the small size of the sample and difficulties in defining the population which it represents, dependable generalizations cannot be made from these data. However, the data would indicate that it was possible to secure the advantages cited above of supplying re-interviewers with the original Census data without serious effects in the direction of "confirming" erroneous entries. As further control for a small proportion of units (5 per cent of the households and 10 per cent of the farms) in the Post-Enumeration Survey sample, no transcription data were supplied the re-interviewer.

CHECKING ACCURACY OF INTERVIEWING

The subsample in which the interviewer was not given the original Census data supplies one check on the accuracy of his work. The relative number and magnitude of the differences when the interviewer is given the original Census data should be about the same as when he does not have the data, unless he has looked at the information instead of obtaining an independent answer.

Another vital problem in the interviewing was making every reasonable effort to insure completeness of the Post-Enumeration Survey canvass. The basic procedure involved giving the interviewer a list of persons, households, and farms which he was to complete. As a check, it was decided to delete the names of some persons actually enumerated in the Census and see whether or not the Post-Enumeration Survey interviewer "picked up" those persons. The re-interviewer's picking-up of the omitted person indicated that he was correctly carrying out the procedure for the detection of missed persons. It is, of course, no guarantee that the procedure was effective in detecting missed persons.

A sample of about 170 Post-Enumeration Survey segments was re-checked by professional personnel of the Bureau to determine the accuracy of the identification of missed households. The results of this re-check indicated that the Post-Enumeration Survey may have underestimated the (net) number of missed dwelling units by 30 per cent. This estimate is, however, subject to a sampling error of 15 per cent. The effect of this deficiency in the Post-Enumeration Survey on the estimates of missed persons appears to be small since the omitted dwelling units were, in large part, vacant units or units with only one occupant.

DESIGN OF THE INTERVIEW

In developing the questions used in the Post-Enumeration Survey, the possibility of simply repeating the original Census questions was

considered and, quite early, rejected. It was felt that this procedure would give a measure primarily of "variance" rather than of accuracy. Every effort was made to incorporate checks on those features which seemed to be a source of error in previous surveys. In many cases, several questions were asked in the Post-Enumeration Survey to elicit data obtained by a single question in the Census. The purpose of multiple questions was to prevent error by calling the attention of respondents and interviewers to any inconsistencies or misunderstandings and to assist the respondent toward a more complete recall of the desired information. The development of the questions on income illustrates this procedure.

In the Census three questions were asked about income. These questions referred to (1) wage and salary income, (2) income from own business, (3) other kinds of income (e.g., pensions, rent, dividends from stocks, bonds, etc.). Considerable experimentation was carried on with question wordings pertaining to income in various Census surveys and also in the pretest of the Post-Enumeration Survey. Most of the variations in the wordings of income questions attempted to check on the specific errors which were thought to be most frequent.

In a special study on housing in Baltimore, a different approach, the "job history" approach, was used to obtain a check on the accuracy of wage and salary income data. The essential feature of this approach is accounting for all periods of employment or unemployment during the year. For each job held, a series of "check" questions on income is asked and for each period of unemployment, questions are asked about unemployment compensation. The job history approach makes for a considerably longer interview but the results of the Baltimore study indicated that the job history approach detects income which may be omitted when other techniques are used. Consequently, for the Post-Enumeration Survey, the job history approach was adopted.

Another important consideration in design of the interview was the specification of the respondent. In the Census, the interviewer was permitted to obtain information from any one of several qualified respondents. It was suggested that the interviews in the Post-Enumeration Survey be obtained from the person himself (except for minor children, mentally incompetent individuals and similar cases). It was anticipated that this would improve the accuracy of the responses (since the person himself is usually the best informed respondent for his own characteristics) but might be excessively costly and might also result in a higher non-interview rate. The costs and non-interview rates were, therefore, explored in the Post-Enumeration Survey pretests. These

pretests showed that the restrictions on the respondent resulted in no significant increase in non-interview rates and somewhere between 40 and 50 per cent additional calls (and, presumably, about this increase in total travel and "contact" costs). It was decided that the cost increase would not be prohibitive and the Post-Enumeration Survey interviewers were told to interview only the person himself (with the exceptions noted). No data are available for determining whether the increased costs were justified. It was proposed that a test of this point be incorporated in the Post-Enumeration Survey but the proposal was rejected as representing a substantial additional complication in both transcription and interviewing.

TIMING OF THE POST-ENUMERATION SURVEY

It was necessary to weigh a possible loss of accuracy if the Post-Enumeration Survey were started late against a loss if it were started early. If the interviewing were to be postponed, respondents would be questioned about an event which might be increasingly remote in time (and possibly in place). This is not serious for "static" items such as date of birth, but is very important for items whose recall tends to be obscured by subsequent events such as jobs held and income received during 1949.

On the other hand, it would have been undesirable to start the Post-Enumeration Survey before the completion of the Census enumeration. Assurance that there would not be a bias in the Census enumeration of units in the Post-Enumeration Survey sample could be achieved only by making sure that these units were not distinguished in any way from other units during the Census. Theoretically, a given segment could have been started as soon as enumeration had been completed for the Enumeration District in which it was located. Actually, since Enumeration Districts, even with a restricted area such as a county, are completed at widely different times, there were two objections:

1. Selection of interviewers from a given county would have been notification that this county was in the Post-Enumeration Survey before all work in the county had been completed.
2. Before the Post-Enumeration Survey work in a segment could start it was necessary to transcribe information from the Census schedules. If this had been done in the local Census office, the local supervisor might have been motivated to have the Post-Enumeration Survey areas checked and the check might have resulted in alterations of the Census results.

Consideration of the dangers of bias led to a decision that transcription be done after all Census enumeration in a county had been completed and sent to the central Census processing sections in Washington and Philadelphia.

In retrospect, the disadvantages of an early start seem as serious as they did at the time; but costs of delay, even beyond the delay anticipated, seem to have been less serious, i.e., observation of the Post-Enumeration Survey interviewing suggests less recall loss and less difficulty in locating respondents than had been feared.

SELECTION OF INTERVIEWERS

It was generally agreed that the nature of the Post-Enumeration Survey required interviewers of top competence. What was meant by "top competence" and how to select such interviewers were matters on which it was almost impossible to secure agreement. Among the qualifications one might list for a "good interviewer" are intelligence (or, more generally, "alertness") and ability to get people to talk freely (secure "rapport"). For Census work, certain less obvious qualifications seem to be important. Coverage is essential and coverage requires meticulousness and persistence; requires, in fact, an "over-conscientiousness" which may be in direct opposition to the characteristics ordinarily associated with a "good interviewer." Checking answers requires an interviewer who will critically evaluate each answer and will not accept it until he has probed the matter to the hilt. However, a critical attitude alone is insufficient—the respondent's cooperation and interest is essential and the interviewer must be able to secure these and still maintain a "cautious skepticism."

All of the above are qualifications hard to define and equally hard to recognize even if found. To set down criteria (with sufficient exactness to make them usable) was practically impossible. One suggested procedure was direct recruitment by persons traveling out of the central office. Such persons might, it was felt, apply adequately criteria of the type cited without elaborate written specifications. Aside from the fact that such an advantage was only hypothetical and the fact that sufficient central personnel were not available and that there are counterbalancing advantages to field selection of the interviewers, considerations of morale in the field organization can be raised against such a procedure. It would seem desirable to let the immediate supervisor have final choice of the interviewers with the accompanying feeling of responsibility for their performance, i.e., not put the supervisor in the position of excusing poor performance as being the work of "a bunch of

deadheads foisted on him by the central office." On the other hand, complete freedom of choice for the local supervisor might, in some cases, mean the placing of personal friendship or other considerations above basic qualifications for the job, particularly if interviewer selection were to be initiated before the supervisor had been given a complete picture of the scope of the job required. These considerations pointed to restricting the supervisor's choice to those candidates meeting specified minimum requirements but, within this class, leaving the final choice in the supervisor's hands.

Another knotty problem was whether supervisors and interviewers should be selected only from persons with previous Census experience. With regard to the supervisors, there was unanimous agreement that previous Census experience was essential, since an inexperienced person would have his energies diverted from his major tasks by difficulties involved in mastering Census administrative procedures, which, however, would be familiar to experienced Census supervisors. It was, therefore, decided to select Post-Enumeration Survey supervisors from experienced members of the Bureau's permanent field staff.

Familiarity with Census concepts and procedures was also desirable in the interviewers and selecting persons with Census experience would mean cost savings in interviewer training. It was also felt that among the 150,000 enumerators employed on the 17th Census, there must be many of outstanding competence and we might be more successful in locating these than in locating competent interviewers among candidates whom we had not had an opportunity to observe in an actual interviewing situation.

As against the cost argument, there was the feeling that restriction of selection to Census enumerators ruled out classes of persons among whom there might be some highly desirable candidates (e.g., graduate students, teachers and professors, and experienced persons from other survey organizations available for the Post-Enumeration Survey during the summer but not for Census work in the spring). It was also felt that the problem of getting reliable reports on an individual's performance as a Census enumerator was considerable since the enumerators worked for only a short time and there was little opportunity to review their work.

Another objection to the use of interviewers who had worked on the Census was the possibility of bias if a person happened to be assigned to check his own work or work done under his supervision. The possible bias arising from an interviewer's checking his own work is trivial, since the probability of cases from any Enumeration District being in the

sample was quite small (and most enumerators worked in only one or two Enumeration Districts). Bias due to the interviewer checking work which he supervised in the original enumeration might be more substantial but could be (and was) avoided by not assigning any interviewer to any area for which he was responsible during the Census.

Since there was no basis for assuming that one method of selection was better than another in terms of quality, it was finally decided on cost grounds to restrict the interviewer selection to persons with experience in the Census. This probably imposed only a relatively minor restriction (since only 260 interviewers were finally selected and the available pool was many thousand). On the other hand, it is by no means certain that another type of selection would have meant either substantially increased cost or lower quality.

It was necessary, in large part, to abandon the attempt to spell out detailed selection criteria. As a first step, selection was restricted to persons who made a high score on the general selection test given to all candidates for enumerators in the regular Census. Additional criteria finally adopted were that (1) the interviewer score in the top two-thirds on a Post-Enumeration Survey selection test which measured ability to read and follow instructions and a schedule of the type actually used (but did not get at any of the personality characteristics cited above); (2) geographic distribution of interviewers be restricted so that about two-thirds of them would live in an area they were to enumerate; and (3) supervisors interview at least 4 candidates for every interviewer position.

TRAINING OF INTERVIEWERS

Considerable experience was available from the Census and other previous studies on techniques of interviewer training. Such techniques had apparently worked satisfactorily and it was decided (mostly as a matter of convenience) to use the same methods with emphasis on:

1. The interviewer going through several actual interviews in the field under observation as part of his training— it was felt that this was important in making the "classroom" training more meaningful.
2. Problem solution and the "role-playing" technique (here, the trainer acts as respondent and is interviewed by one of the trainees while all of the trainees record the answers).

It had been decided to use as interviewers only persons with previous Census experience. However, consideration of the possibilities of inade-

quate original training and development of "bad habits" during their previous work, led to the decision that the Post-Enumeration Survey training repeat and (on basic points) amplify the Census training. The Post-Enumeration Survey interviewers were given considerable additional material on maps and canvassing (which were felt to be important for good coverage) and on interviewing techniques, particularly the problems of verifying an answer and probing for reasons for discrepancies from the original Census response. Considerable emphasis was placed upon explaining to the interviewers the reasoning behind the phrasing of the questions and the design of the procedures.

In training the supervisors they were first put through the training designed for the interviewers (combining, with appropriate modifications, the two training courses planned for rural and urban interviewers). Examination of this experience indicated that training of rural interviewers was likely to prove unsatisfactory unless the material was simplified or the length of the training extended. To avoid a longer training period, the procedures planned for the rural interviewing were reviewed. This called attention to the fact that in checking population and housing characteristics a rural interviewer was required to use the same questionnaires used by an urban interviewer. The rural interviewer filled out these questionnaires only for missed persons and missed dwelling units and, for this purpose, it was not necessary that the information be any more precise or detailed than that collected for enumerated units in the original Census. Actually, this rural procedure was a carry-over from the earlier planning before it was decided to split the interviewer's job. Although time was short and the training and interviewers' manuals were already printed, it was decided that the cost of revision would be less than the cost (and difficulties) of carrying out the original plans. A Census schedule which had been used for self-enumeration was, therefore, revised by adding questions regarding the usual residence at the time of the Census of presumptively missed persons and this schedule was adopted for use by the rural enumerators. Although the work of revising the schedule and manuals was considerable (and the timing was hectic), the revision permitted simplification of the rural interviewer's job and of his training.

MODIFICATION OF THE SURVEY DESIGN

Having made the above decisions, work was started on drawing the sample, preparing the necessary maps and control lists, designing and printing of questionnaires and drafting of instructions. However, the design adopted was to be viewed as only an initial working document.

While the survey could have been carried through as initially designed, facts brought out in the course of the work dictated important modifications. It was, in fact, a conscious policy to maintain flexibility of plans up to the point where efficiency required that they be frozen.

As indicated above, the sample was intended to yield segments of approximately 6 households per segment in urban areas and 10 households per segment in rural areas. Figures for about the first 20 per cent of the segments drawn indicated an average size considerably larger than the expected 7.6 households per segment. The complete sample, before the subsegmenting described below, had an average size of about 35 households per segment. About 10 per cent of the segments had over 35 households with an average of about 300 households. This was due to the fact that, in many cases, although there was known to be a large number of households in an area, the available maps did not show the detail necessary for subdividing the area and it had to be selected as a whole (with a lower weight). In most surveys, this difficulty can be overcome by subsampling within the segment and interviewing the units in the subsample. For checking coverage, however, it was desirable to have every household in a segment interviewed. One solution would have been to have the Post-Enumeration Survey interviewer divide the larger segments at the time he visited them and cover only one randomly selected subsegment. The difficulties of getting the Post-Enumeration Survey interviewer to execute this task satisfactorily, along with his other duties, made this solution unacceptable. Another alternative was to do the subsegmenting as a separate operation with the Post-Enumeration Survey interviewers, but before their training for the actual interviewing. This was rejected because it meant recruiting the interviewers earlier than had been planned (which was undesirable both because it meant disclosing prematurely some of the areas included in the Post-Enumeration Survey and because it meant a very awkward timing). The solution finally adopted was to do the subsegmenting with personnel from the Bureau's Geography Division, a solution which avoided the problem of using the Post-Enumeration Survey interviewer and had the additional advantage of doing the job with exceptionally well-qualified persons.

As the sample was drawn and detailed specifications of the procedure were developed, more precise estimates of total costs could be made. It became evident that total costs would exceed the budgeted amount unless the design was modified. Since the purpose of the Post-Enumeration Survey was to check the Census results, there was a strong commit-

ment to obtain the maximum accuracy on a given interview even if this meant a relatively high cost per case. This ruled out effecting economies at the risk of reducing quality (e.g., by permitting an interview with any responsible person instead of requiring an interview with the person himself). The alternative was to reduce the sample size. The most obvious way to cut costs was to eliminate some primary sampling units from the sample. The objection to this procedure was that it meant either completely re-drawing the sample (scrapping the work done to date) or ending up with strata of very unequal size which would have increased the sampling error disproportionately to the cost savings. However, oversampling of the primary sampling units in the West (introduced originally to permit separate estimates for this region for agricultural characteristics) could be eliminated without appreciably affecting the variance of national estimates. This decision was adopted and reduced the number of primary sampling units to 270.

As originally designed the sample called for interviews with all persons in the sample households. While this procedure is relatively simple, it is not necessarily efficient. Furthermore, on some characteristics, (e.g., immigration status, education and income) the Census had obtained information only for every fifth person so that for these characteristics the Post-Enumeration Survey sample would have been one-fifth the size of the sample for other characteristics (e.g., age, birthplace, occupation). In view of the budget situation, it was felt that cutting the sample for those characteristics obtained of every person in the household would not be too serious. Therefore, those persons for whom the Census asked all questions and one-quarter of the remaining persons in the Post-Enumeration Survey sample households were retained in the sample.

Another economy was effected by permitting Post-Enumeration Survey interviewers to obtain information for missed units from any responsible respondent. The aim of the restrictions on the respondent for enumerated units was to improve on the original Census data. With respect to the detailed information obtained for missed units, it was only necessary to attain the accuracy level required in the original Census.

RECORD CHECKS

The design of the Post-Enumeration Survey emphasized securing results of high accuracy. Nevertheless questions were raised about the validity of the Post-Enumeration Survey: How can one be sure this

technique would give a better answer? How can you check the errors in one survey by taking another? This line of questioning frequently ended with a suggestion that there were records (presumably authoritative) which would give a better check than a re-interview.

However, almost all record-matching studies end with unmatched cases and cases where the match is questionable. The errors found for the matched cases do not provide an estimate for all cases unless one is willing to make arbitrary assumptions with regard to the errors among unmatched cases. Even where a match has been established, one cannot conclude that the records are necessarily accurate. Where a birth certificate was filed soon after the birth, the date recorded on it is, in general, dependable. Where it was filed several years later, one might have considerably less confidence in the recorded birth date. The accuracy of record data varies from one source to another and also from item to item on the same record. A record in the Veteran's Administration files may provide conclusive evidence that John Doe is a veteran but not that he was born on the date stated on that particular record.

The above considerations sharply limit the utility of record checks. Where errors are large for the matched cases, however, a record check provides a valuable warning of unreliability of data and useful leads for further investigation of sources of error.

It was decided to experiment with record checks for age, birthplace, citizenship status, highest grade of school completed, income, veteran status, and industry in which employed. In pretests of the Post-Enumeration Survey, additional data were included which might be useful in checking these items, e.g., an inquiry regarding branch of service and serial number for use in checking veteran status. In completing the matches for the pretest data, our attention was called to additional items which would improve the efficiency and the accuracy of the match. These items were added to the Post-Enumeration Survey questionnaire. After collection of the data, the check of highest school grade completed was dropped on the grounds that the effort involved would have been out of proportion to the expected success.

IMPLICATIONS FOR FUTURE RESEARCH

Although the Post-Enumeration Survey possessed certain distinctive features peculiar to its purposes and content, its history illustrates the complex decision process involved in survey planning. This decision process requires the balancing of cost and accuracy considerations both within a particular aspect of a survey and among different aspects. The

over-all cost of a survey and the accuracy of the estimates the survey provides are resultants of the costs and accuracies associated with a variety of processes—sample selection, interviewing, coding, tabulating, etc. Since most surveys have a fixed over-all expenditure limit, increased expense to improve the accuracy of one phase will usually require curtailed expenditure and lowered accuracy in some other phase—for example, may require a decrease in sample size with consequent increase in variance. Between survey processes a balance is necessary; to achieve a reduction of 5 per cent in coding error at the expense of decreased checking of the interviewing with an attendant increase of 10 per cent in interviewing error is hardly desirable, particularly if the interviewing error is already larger than the coding error. Within a given survey process—for example, interviewing—an increased expenditure for training must be balanced against an increase in expenditures for supervision, interviewer salaries, more detailed questions, etc., in terms of the relative effect of the various expenditures on the quality of the interview output.

In the process of balancing errors and costs between different survey operations, it is important to distinguish between the "variable" errors which decrease with increasing size of sample and the biases—which are unaffected by the sample size. A survey design may be characterized by its sampling variance and bias and, in addition, by the variance and bias arising from other sources, such as interviewer, respondent, and processing error. By "nonsampling variance" is meant the variability of errors around the average error which is the bias. Depending upon the relationship between variable errors and biases, increased expenditure per case may or may not lead to a decrease in the over-all error. For example: Suppose that Method A with an expenditure of $1.00 per case would, if applied to the whole population, give an error (bias) of .2 year in an estimate of average schooling and if the estimate were based on a random sample of a single case, give an (expected) error of 10 years. Assume, also, that Method B with an expenditure of $20 per case would give an error (bias) of .05 year if applied to the whole population and of 8 years if applied to a single (randomly selected) case. The difference between the two methods arises, let us say, from greater expenditures for training and checking on all survey processes in the case of Method B. If cost were not a factor, Method B would be preferable. In most cases, cost is a factor and operates differently depending upon the total funds available; using a simple random sample, Method A would give a mean square error of .060 for a total survey budget of

$5,000 compared with a mean square error of .258 for Method B. With a total survey budget of $100,000, the relationship would be reversed, Method A showing .041 mean square error and Method B .015.[1]

In planning a survey, it is not sufficient to know that one technique gives more accurate results for individual cases than does another technique. The amount of the improvement and the relative costs of the two techniques must also be known. In addition, the relationship among the different phases of a survey may be extremely important. Since the error in a survey estimate is not a sum of the errors in the different survey operations but tends rather to be the square root of a sum of squares of errors, reduction of a given amount in a relatively large source of error has more impact on total error than an equal reduction in a small source. For example, reducing the response bias in an estimate of average income from $250 to $200 will have more effect on total error than reducing the sampling error (standard deviation of the estimate) from $100 to $50.

Reasonably satisfactory techniques are available for assessing the gain of a given sampling design in terms of its effect upon cost and accuracy, and for evaluation of processing operations. At present, techniques available for the assessment of a given interviewing procedure are not nearly so satisfactory. In some cases there may be evidence that one interviewing method gives greater accuracy than some other method. However, in almost all cases, we have had no specific determination of *how much* improvement in accuracy is obtained by a given interviewing technique and cannot, therefore, determine whether the gain in accuracy justifies the increase in cost. Although present methods for measurement of the error of an interview may fall far short of perfection in giving an objective evaluation of quality, they can often provide relative measures of reliability in the form of lower bounds. Such data are useful and usable in improving survey design. They also will provide a basis for improving techniques of measurement. The development of more satisfactory techniques for the measurement of interview error, and collection of data for the evaluation of specific survey methods, constitute in our opinion the most essential steps to be taken toward the improvement in survey design.

---

[1] M.S.E. for Method A $= (.2)^2 + \frac{(10)^2 - (.2)^2}{n}$

M.S.E. for Method B $= (.05)^2 + \frac{(8)^2 - (.05)^2}{n}$

n = sample size for specified budget

REFERENCES

Bancroft, Gertrude, and Welch, E. H., "Recent experience with problems of labor force measurement," *Journal of the American Statistical Association*, 40 (1945), 205–13.

Beales, LaVerne, "Negro enumeration of 1920—a reply to Dr. Kelly Miller," *Scientific Monthly*, XIV: 4 (April 1922).

Carson, Daniel, "Incomparabilities in Census of Manufactures data," *American Economic Review*, XXI: 3 (1941), 559–60.

Davis, I. G., "A discussion of the accuracy of agriculture census enumeration in the Northwest," *Journal of the American Statistical Association*, 28 (1933), 272–86.

Ducoff, Louis, and Bancroft, Gertrude, "Experiment in the measurement of unpaid family labor in agriculture," *Journal of the American Statistical Association*, 40 (1945), 205–13.

Miller, Kelly, "Enumeration errors in Negro population," *Scientific Monthly*, XIV:2 (February 1922).

Myers, Howard B., and Webb, John N., "Another census of unemployment," *American Journal of Sociology*, 60 (1937), 521–33.

Myers, R. J., "Errors and bias in reporting of ages in census data," *Transactions of the Actuarial Society of America*, XLI, Part 2, No. 104 (October–November, 1940).

Myers, R. J., "Under-enumeration in the census as indicated by Selective Service data," *American Sociological Review*, 13:3 (1948), 320–25.

Palmer, Gladys L., "Factors in the variability of response in enumerative surveys," *Journal of the American Statistical Association*, 38 (1943), 143–52.

Price, Daniel O., "A check on under-enumeration in the 1940 Census," *American Sociological Review*, 12:1 (1947), 44–50.

Van Kleek, Mary, "The federal unemployment census of 1930," *Journal of the American Statistical Association*, 26 (1931), 189–200.

ON THE DISTINCTION BETWEEN ENUMERATIVE AND ANALYTIC SURVEYS*

W. Edwards Deming
Bureau of the Budget and New York University

I. DEFINITION OF THE ENUMERATIVE AND ANALYTIC USES OF DATA

Purpose of the paper. Statistical data are supposedly collected to provide a rational basis for action. The action may call for the enumerative interpretation of the data, or it may call for the analytic interpretation.

The aim here is to exhibit some of the consequences of failing to distinguish between the enumerative and the analytic uses of data. This distinction is necessary in the statement of the aims of a survey, census, or experiment, in order that the plans for the collection of the data and for the tabulations may most economically meet the needs of the consumer, and it is equally important in the interpretation of data.

Thus, to draw on a result from a later paragraph, information obtained in a complete census concerning every person in an area (e.g., on occupation, income, or education) still possesses *for analytic purposes* a sampling error that is actually about a quarter as great as the sampling error of a 6 per cent sample. The consequences are far-reaching. In using a census-table for analytic purposes, even though the figures come from a perfect complete count, it is therefore necessary to bear in mind that small numbers in a cell are unreliable in the sense that they have a standard error, just as if they had arisen in sampling, as indeed they did. Moreover, in the planning of a complete census, it is therefore imperative to use sampling for every bit of information that is not necessary as an aid to complete coverage, or required to give detail for small areas (such as the block statistics). Name, relationship to the head, age, sex, marital status, color are probably all necessary for the sake of completeness of coverage. These things, plus a few questions on rent, tenure, year built, will provide the information required for the block statistics.

To draw on another result, we shall see that it is often impossible to design a survey that will supply economically information for both enumerative and analytic purposes. For example, in a marketing survey, the best design for an estimate of the number of people who prefer to use ground coffee at home, rather than soluble coffee, requires, for

* Delivered at a conference on sampling conducted by the Institute of Statistics, University of North Carolina at the Blue Ridge Assembly, 21 July 1952.

greatest economy, one type of sample design; whereas a study of the *reasons*, or even of the difference in the two proportions, requires another design. One must be prepared to make some sacrifices in precision, as it may not be economical to satisfy both aims simultaneously.

The distinction between the enumerative and analytic uses of data.[1] Briefly, the enumerative question is how many? The analytic question is *why?* is there any difference between the two classes, and if so, how big are the differences?

In the enumerative problem, some action is to be taken because the frequency of some particular characteristic of the universe is found to exceed some critical value. The crop of wheat, according to a sample survey, turns out to be large or small. As a consequence, the market goes down or up, and production of meat, cereal products, and of substitutes shifts one way or another because of this information. The Census, or perhaps a sample study of birth registrations, shows that in a particular city the number of children in the primary schools will be much greater in 4 years than now. Bonds are issued; work commences on a new school building. Inspection of a sample of wool or of cotton may determine its disposition, the price to be paid for it, and what kind of cloth and of garments to make of it. Inspection of a lot of industrial product determines whether it will be accepted or subjected to screening or to a lower rating, or outright rejection.

Such problems are enumerative because they depend purely on a determination of the number of people in an area, or the inventory of grain, or the production of grain, or the quality of a product. They do not involve the analytic question of *why* all these people are there or why the crop this year is what it is; or *why* the wool or cotton or product is so good or so bad.

When certain cities in America were swelled with in-migrants because of war production in the spring of 1944, special censuses were taken with the aim of arriving at equitable allocations of food, gasoline, repair parts for buses and trolley cars, and other necessaries of living. Equitable distribution of supplies to these cities was impossible because no one knew just how many people were in them: assertions of editors and chambers of commerce did not provide a basis for action. The problem was enumerative because the action (viz., allocation

[1] These are the terms that I invented for Chapter 7 in my book *Some Theory of Sampling* (John Wiley, 1950). The terms are not important; the concepts are. The concepts are old, but plain statements of what they are and of the consequences of failing to keep them in mind in design and analysis are not easy to find. Similar but not exactly parallel concepts occur in the analysis of variance, under the terms Model I and Model II, a lucid explanation of which occurs in the paper by Churchill Eisenhart, "The assumptions underlying the analysis of variance," *Biometrics*, 3 (1947), 1–21.

of food and materials) depended on how many people were there and not why they were there.

By law, the Social Security program is partly an enumerative problem because federal reimbursement to a state depends on the number of inhabitants 65 and over within the state. Public health programs, agricultural adjustments, and other allotments depend on population and acreage, and are examples of enumerative uses of data. Administrative problems concerned with the long-range aims of these programs, however, are analytic.

In the analytic problem, the action is to be directed at the underlying causes that have made the frequencies of the various classes of the population what they are, in order to govern the frequencies of these classes in time to come. Familiar examples of analytic studies are found in intelligent city-planning. More familiar studies are the differential effects of varieties and of treatments in agriculture and entomology. The particular crops that are measured are of interest only because they aid decisions on what varieties and treatments to use for the best results in crops *yet to be planted*. We may run an experiment with a group of test animals or with patients in a hospital, but when we generalize from these tests we are thinking of the production process: what will it produce in the future? The present tests are important only because they help us to prescribe or to modify the treatment for future use. The control-chart is a splendid example, the purpose being to control the production process and the quality of lots yet to be made. Other examples are medical and social studies wherein interest centers in the causes that produce differences in health, fertility, or death-rate in different segments of a population of people. Current population surveys in the United States and Canada aid studies of employment, unemployment, farm and industrial labor, school attendance, etc. The monthly sample of deaths by causes, published by the National Office of Vital Statistics in the United States, aids in the control of epidemics and the spread of disease. Its use is both enumerative and analytic.

Special reference to the statistical control of quality. Both the enumerative and analytic problems present themselves hourly in the statistical control of quality. A batch of product has been produced, let us suppose, and the machine is already producing another batch. Two questions arise: B (analytic). Shall we leave the machine alone, or shall we adjust it? Shall we make it run slower or faster, or shall we change the type of chemical bath? A (enumerative). What shall we do with

the batch of product just made? Shall we send it on to the next operation (which might be into the consumer's market, or into the next bay of the same factory for further work)? Or is the product so defective that we must re-work it, sell it as second-class, or scrap it?

A chemical engineer whose specialty is the production process may have a special interest in Problem B, and little in Problem A, which he leaves to someone else. On the other hand, if we are the purchaser of a batch of product, such as a single automobile, or some paint for our home, or a carpet, we certainly have a special interest in Problem A. We wish to know the quality of this particular batch of product. It is little comfort to know that the process by which it was made was a good one, and was in a fine state of statistical control, if the product that we ourselves purchase turns out to be defective and unsuited to our purpose. A manufacturer, on the other hand, must purchase raw materials and assemblies in quantities, week in and week out. In order to cut the costs of these materials and to improve their quality, he must concern himself not only with Problem A, the inspection of these materials upon receipt; he must in addition take a lively interest in Problem B, the control of the production processes in the plants of his suppliers.

The methods of the Shewhart control chart are essentially analytic, as they tell when to take action on the process. In contrast, the methods of acceptance sampling are primarily enumerative, dealing with the disposal of a lot, although they react secondarily on the process by forcing better control where needed.

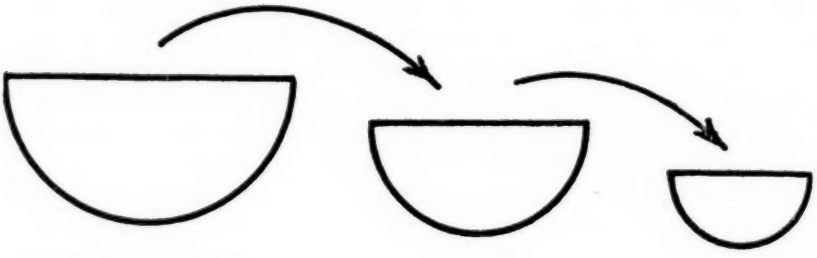
II. THE SAMPLING VARIANCES FOR THE TWO TYPES OF DISTRIBUTION

The two uses re-stated in terms of sampling distributions. Re-stated in terms of a mechanism for carrying out the sampling, we may distinguish between the two uses (enumerative and analytic) by consideration of the two distinct types of repetition of the operations that lead to two distinct sampling distributions. In the enumerative case, we take repeated random samples from the same lot, and seek the sampling distribution of the mean or of other statistical measures of these samples. In the analytic case, we take repeated random lots from a supply or cause system, and we select a random sample from each lot; then seek the sampling distribution of the mean or of other statistical measures of these samples.

The use to which the data will be put determines which of the two

types of repetition is applicable in any one problem. Unfortunately, sometimes we require data from the same survey to serve both purposes.

It is helpful to look at a diagram. The figure shows three bowls with poker chips, all physically similar, some red and some white. By stirring the contents of any bowl thoroughly, and reaching in blind-folded, it is possible to satisfy satisfactorily the conditions for a random sample. Another way is to give serial numbers to the chips and to draw



Cause System or Production Process	Lot	Sample
Initially	*Initially*	
Mp red	NP red	r red
Mq white	NQ white	$n-r$ white
M total	N total	n total

them with random numbers. The bowl on the left represents the process or cause system. It is a supply of chips. The bowl in the middle represents the lot. It is the people in an area today, or a batch of product, or a crop. The lot has come, we suppose here, as a random sample from the process. This assumption is over-simplified; nevertheless it is a first step to an understanding of analytic problems. The small bowl at the right represents a sample drawn from the lot.

The four possible different variances. Now we are able to state the two problems in terms of estimation. In the enumerative problem the sample is used for an estimate of the contents of the lot, which is described by the proportions P and Q. In the analytic problem the sample is used for an estimate of the contents of the supply, which is described by the proportions p and q. The same sample serves both purposes, but not equally well, for the two estimates have different variances. Hence the proper size of sample, and how to interpret a sample, will depend on whether the aim is enumerative or analytic.

The variances of r/n as estimates of p and of P are shown in the accompanying table. There are four cases, A, B, C, D, depending on how the lot and the sample are drawn.[2] There are two interpretations (enumerative and analytic) in each case.

Table of the variances of $\hat{p}$ and of $\hat{P}$
In all cases, $E\hat{p}=p$ and $E\hat{P}=P$

The N balls in the lot-container are drawn from the supply	The sample of n is drawn from the lot-container	
	With replacement	Without replacement
	Case A	Case B
Analytic	$\text{Var}\,\hat{p}=\frac{pq}{n}\left\{1+\frac{n-1}{N}\right\}$	$\text{Var}\,\hat{p}=\frac{pq}{n}$
WITH REPLACEMENT		
Enumerative	$\text{Var}\,\hat{P}=\frac{PQ}{n}$	$\text{Var}\,\hat{P}=\frac{N-n}{N-1}\,\frac{PQ}{n}$
	Case C	Case D
Analytic	$\text{Var}\,\hat{p}=\frac{pq}{n}\left\{1+\frac{n-1}{N}\,\frac{M-N}{M-1}\right\}$	$\text{Var}\,\hat{p}=\frac{M-n}{M-1}\,\frac{pq}{n}$
WITH REPLACEMENT		
Enumerative	$\text{Var}\,\hat{P}=\frac{PQ}{n}$	$\text{Var}\,\hat{P}=\frac{N-n}{N-1}\,\frac{PQ}{n}$

APPLICATIONS

Tabulation plans. The two variances, $\text{Var}\,\hat{p}$ and $\text{Var}\,\hat{P}$, are different. A sample therefore contains different amounts of information for the two purposes. How do these observations help us in the design of samples? They tell us that if there is a definite enumerative aim in finding out how many people there are with a given characteristic however rare, then the tabulation and printing of small cells may be justified, provided the universe will still retain enough of its characteristics by the time the sample is tabulated.

On the other hand, if the aim is analytic, there is the sampling error $\sigma_{\hat{p}}$ even if the figure comes from a complete census, and this error may become troublesome in small cells. It will then be well to economize by

[2] For the derivations of the variances see the author's *Some Theory of Sampling* (John Wiley, 1950), p. 254.

using sampling in the collection of such data, and to determine in advance what consolidations may be made in the tabulating and in the printing. Much space and money are wasted annually on the tabulation of cells that are too small for analytic use, and which have no enumerative use. Too often the excuse for tabulating small cells in a complete census is that they came from a complete census and must therefore be correct. When the use of such tables is analytic only, such arguments do not hold: a reconsideration is due.

Case B is one that often corresponds approximately to many problems in real life. For enumerative use we take the proportion r/n in the sample to estimate the proportion P in the lot. The variance of the estimate

$$\hat{P} = \frac{r}{n} \tag{1}$$

is seen in the table to be

$$\text{Var } \hat{P} = \frac{N-n}{N-1}\,\frac{PQ}{n}, \tag{2}$$

which reduces to 0 if the sample n is increased to a complete census of the lot, when $n = N$.

In contrast, for analytic purposes in Case B we use the proportion r/n in the sample to estimate the proportion p in the supply. The variance of the estimate

$$\hat{p} = \frac{r}{n} \tag{3}$$

is seen in the table to be

$$\text{Var } \hat{P} = \frac{pq}{n}. \tag{4}$$

The size N of the lots, although they furnished the samples, does not enter into this variance at all. The size N only limits the size of the sample: it cannot be bigger than the lot. To reach greater precision than pq/N (the variance of a complete census) we must draw another lot and sample it also; then combine the two estimates of p.

Effectiveness of a medical treatment. We may see from the following example the two possible ways of interpreting data. Three hundred ninety-eight patients in a hospital went under treatment for a certain

disease. After two months a random sample of 250 patients drawn from the list of the original 398 showed no sign of recurrence. How do we figure the odds that there are 0, 1, 2, etc., recurrences in the remaining 148 patients? Stated another way, what is the highest number of recurrences in the original 398 that would permit such a result as often as once in 20 repetitions of the sample? This question is framed as an enumerative one.

This sample was drawn without replacement, and we may see from either Case B or D that the probabilities to apply are the terms of the hypergeometric series.

Let K be the number of patients in the original 398 who actually have a recurrence, none of which showed up in the sample of 250. Then the probability of the result obtained is given by the hypergeometric term

$$P_e = \frac{\binom{398-K}{250}\binom{K}{0}}{\binom{398}{250}}$$

$$= \frac{398-K}{398}\,\frac{398-K-1}{397}\,\frac{398-K-2}{396}\text{ etc. to 250 fractions.} \quad (5)$$

A few results are tabled here, whence we may conclude with odds of about 19:1 that there are not more than 3 recurrences in the lot.

$K=NP$	P	P_e
0	0	1
3	.008	0.051
4	.010	.019
5	.012	.007
6	.015	.002
7	.018	.001

Now suppose that the problem is to predict the proportion that would be cured in a *succession of lots* of patients. This is an analytic question, and the theory to use here then is the binomial series. If p is the proportion of recurrences in the general population (from which by assumption now the 398 patients is a random sample), then the probability of the observed result is

$$P_e = q^{250}p^0 = q^{250}. \tag{6}$$

It is interesting to note that the size 398 of the lot does not come into this probability at all.

A few results are in the table here, whence by interpolation we may conclude with odds of about 19:1 that the proportion p of recurrences in the general population would not be more than 12 in 1000.

p	P_e
0	1
.005	.286
.010	.081
.015	.023
.020	.006
.025	.0002

This example, though oversimplified, may help to guide the design and interpretation of the results of samples. It shows specifically how the interpretation changes when we change our aims from the enumerative to the analytic use.

Allocation of sample. In the symbolism just introduced, the analytic aim is to measure the difference between the two proportions p_1 and p_2 which exist in two cause systems, or to find out if there is any significant difference between p_1 and p_2. We cannot examine the cause systems directly; we can only study two groups of farms, plots, patients, or pupils that the two cause systems have produced. That is, we can only study two *lots*, one from one cause system, and one from another. We shall assume that Case B fits the actual events.

The lots may be of different sizes, N_1 and N_2. At any rate, we take samples therefrom of sizes n_1 and n_2, and we ask what should their sizes be to minimize $Var\ (\hat{p}_2 - \hat{p}_1)$, an analytic purpose. The optimum allocation for this purpose requires that

$$n_1 = k\sigma_1 = k\sqrt{p_1q_1} \tag{7}$$

$$n_2 = k\sigma_2 = k\sqrt{p_2q_2} \tag{8}$$

where

$$k = n/(\sigma_1 + \sigma_2). \tag{9}$$

Now usually, if not always, such problems require the aid of statistical techniques only if p_1 and p_2 are not far apart; if the difference between

them were wide, we could observe it without help. Hence we may properly take

$$n_1 = n_2$$

as optimum. Thus, the usual practice of taking equal sizes of sample for clinical or laboratory tests is correct for minimizing the variance of the estimated difference, regardless of the sizes of the populations or of the acreages of the crops whence the samples were drawn.

In contrast, for the enumerative purpose of estimating the over-all average $\bar{x}$ or total X of some characteristic (average rent, total number of unemployed, total acreage in wheat) in the two lots of size N_1 and N_2, the aim is to minimize the variance of $\bar{x}$ or of X, an enumerative purpose. To this end, the optimum size of sample will be

$$\left.\begin{aligned} n_1 &= nN_1/N \\ n_2 &= nN_2/N \end{aligned}\right\} \quad \text{[for proportionate allocation]} \tag{10}$$

or, if one prefers,

$$\left.\begin{aligned} n_1 &= kN_1\sigma_1 \\ n_2 &= kN_2\sigma_2 \end{aligned}\right\} \quad \text{[for disproportionate allocation]} \tag{11}$$

where

$$k = n/(N_1\sigma_1 + N_2\sigma_2).$$

Obviously, the optimum allocations for the two purposes will be different except when the two lots N_1 and N_2 are nearly equal in size.

Unfortunately, the purposes of a survey are often both analytic and enumerative. In a survey to assist the marketing of a certain brand of frozen orange juice, we need to know not only how many people of various income levels buy frozen orange juice of a particular brand, but of all brands, and tinned unfrozen juices as well, and probably fresh fruit besides. These are *enumerative* counts. Then also, probably more important, the survey must discover *why* people of various groups buy or do not buy frozen orange juice and the products that compete with it. This kind of question is *analytic*. The design that is economical for one type may not be economical for the other.

In another study, a research worker wishes to study the variation in the behavior of people, classified by age, education, marital status; perhaps also by religion, urban and rural residence, and occupation. Frequencies are important; but so also are the contributing causes. Thus, this research presents also both enumerative and analytic problems.

One way out is to conduct two surveys—one for enumerative purposes to ascertain the frequencies of certain behavior by class; another, with a more intensive questionnaire, to study the causes.

Another solution is to make some sort of compromise in the design, sacrificing economy and precision in (e.g.) the enumerative results, in order to gain something for the analytic uses. How much to sacrifice, how far to lean, and which way, can be settled only with consideration of the risks and of the losses of making a wrong decision on the basis of information not sufficiently precise, and on consideration of the additional cost of getting more precise information.

A note on acceptance sampling. The probabilities that one encounters in the analytic problem in Case B justify the customary 3-sigma control limits in the form

$$\bar{x} \pm \frac{3\overline{R}}{d_2\sqrt{n}} \quad (\overline{R} \text{ is the average range over a series of samples, and } \overline{R}/d_2 \text{ is an estimate of } \sigma)$$

for the $\bar{x}$-chart, or

$$\bar{p} \pm 3\sqrt{\frac{\bar{p}\bar{q}}{n}} \quad (\bar{p} \text{ is the average fraction defective over a series of samples})$$

for the p-chart, as an aid for detecting uncontrolled variability. It will be observed that the size N of the lot does not appear in these equations even if the sample-size n is 100 per cent of N. This form of computation is now seen to be correct; it is not an approximation. The justification is the absence of N and of any finite multiplier in the analytic Case B.

On the other hand, many writers in dealing with the *producer's risk* in acceptance sampling (the probability that a lot of acceptable quality. will be rejected) have recommended hypergeometric terms (like that in Eq. 5), or rather have reluctantly used binomial terms as approximations to the hypergeometric terms. Actually, however, the producer is concerned with the problem of keeping his process in control and at a desired level p. The quality P of the lots that he produces will vary from lot to lot, yet the risk (probability) that a lot will be rejected on a single-sampling plan turns out to be a sum of binomial terms typified by

$$\binom{n}{r} q^{n-r} p^r,$$

into which the size N of the lot does not enter at all. This problem be-

haves like the *analytic* Case B, for which the binomial terms are correct; they are not approximations.

The *consumer's risk* is another story. The consumer is concerned with the particular lots of product that he is purchasing, and he has a sample from each lot on the basis of which to decide whether to accept or to reject the lot. He may aim to guard against accepting lots with too high a value of P, regardless of p and of the state of control. To compute on this basis the correct probabilities for the consumer's risk, one requires hypergeometric terms. The finite multiplier $(N-n)/(N-1)$ then appears in the variance of $\hat{P}$, because this is the enumerative Case B.

There are thus, strictly, two operating-characteristic (O.C.) curves, one for the producer, another for the consumer. In practice, however, except when the sample is 20% or more of the lot, the two curves coincide, almost, fortunately, and one curve suffices.

The distinction made here between the different probabilities for the producer's and consumer's risks is not new. It forms the basis for the Dodge-Romig tables, as is clear from their text. An extremely lucid exposition appears in the book *Sampling Inspection* by the Statistical Research Group, Columbia University (McGraw-Hill, 1947), pages 183 and 184. I quote:

> There are two alternative ways of interpreting "percentage of defective items in submitted product," and these lead to somewhat different O.C. curves for small inspection lots. (a) The percentage of defective items can be considered as applying to each inspection lot separately. . . . (b) The percentage of defective items can be considered as applying to the process. . . . If interpretation (b) is adopted, the resultant O.C. curve does not depend on the size of the inspection lot.

CONFIDENCE INTERVALS FOR THE NUMBER SHOWING A CERTAIN CHARACTERISTIC IN A POPULATION WHEN SAMPLING IS WITHOUT REPLACEMENT

LEO KATZ
Michigan State College

I. SUMMARY

IN A number of different sampling situations, we encounter a problem which may be stated as follows: From a population of N objects, each of which possesses or does not possess a certain characteristic, we select *without replacement* a random sample of n objects. Observing that m of these possess the characteristic in question, we wish to estimate the total number M having the property. This problem arises in sample surveys of human populations and in sampling inspection of manufactured articles as well as in other, less obvious, situations.

It is well known (see e.g., F. F. Stephan's solution, quoted by Deming [2, p. 294]) that the maximum likelihood estimate for M is the largest integer in $(m/n)(N+1)$. The problem of confidence interval estimation for M, however, seems to have escaped systematic investigation in the literature. In this paper, the problem is considered and two approximate methods are given for construction of these confidence intervals. On the basis of some preliminary investigations, it appears that the second approximation, which allows for correction for discontinuity in the observations, agrees quite closely with the exact computations.

II. FORMULATION OF THE PROBLEM

The roughly similar problem of confidence interval construction for the parameter π in an infinite population or where sampling is done with replacement has been exhaustively studied by Clopper and Pearson [1]. When N is very large, confidence intervals for $M = N\pi$ may be obtained in terms of the well-known results for π. When, however, N is small or moderately large, this approximation is not useful and exact methods or better approximations are needed.

Consider the statement of the problem as one involving a fourfold table. There are two groups of n and $(N-n)$, respectively. We observe that m of the n possess a certain characteristic. Assuming the grouping to be arbitrary and random, we wish to estimate the total number, M, in the combined groups, possessing the characteristic. Accordingly, we have the fourfold table,

$$\begin{array}{c|c||c} m & n-m & n \\ \hline M-m & N-M-n+m & N-n \\ \hline\hline M & N-M & N \end{array},$$

where m, n and N are known at the outset, and M is to be estimated. It is well known, due to a result of R. A. Fisher given by Yates [6], that the exact distribution of m is hypergeometric with probabilities given by

$$(1) \quad P_r\{m \mid N, n; M\} = \frac{n!(N-n)!M!(N-M)!}{N!m!(n-m)!(M-m)!(N-M-n+m)!}$$

where the semicolon before M emphasizes that N and n are known numbers while M is a fixed but *unknown* parameter in our problem.

The known exact distribution (1) defines implicitly confidence intervals for M in terms of the fixed values of n and N and the observed value of m. The explicit construction of confidence intervals requires the existence of tables of the distributions; it is at this point that the difficulties occur. Finney [3] gives percentage points of the distributions for n and $(N-n)$ up to 15 each. The author [5], in unpublished tables, has computed the cumulative distribution functions for all combinations with N up to 24. Thus, except for extremely small numbers, the exact confidence intervals are not available. We turn, therefore, to approximations.

Chung and DeLury [7] have recently published a set of charts[1] giving 90 per cent, 95 per cent and 99 per cent confidence intervals for $N=500$, 2500 and 10,000 with sampling rates of .05 and .1 by tenths to .9. Interpolation for other population sizes using $1/\sqrt{N}$ as argument is satisfactory; interpolation for other sampling rates than those specified is inconvenient, because neighboring values are read from different charts. Extrapolation beyond $N=10{,}000$ is reasonably accurate.

This leaves a gap between about $N=25$ and $N=500$, not likely to be filled by exact computation. The approximation given here may be useful in bridging this gap. For example, it might be possible to obtain approximate hypersurfaces for upper and lower confidence limits by this means, and then to devise systematic corrections in the direction of relatively few exactly computed points.

---

[1] See notice in this *Journal*, 46 (1951), 394–5.

III. APPROXIMATE CONFIDENCE INTERVAL CONSTRUCTION

In the paper by Yates [6] to which reference was made above, it was shown that the exact distribution of m may be approximated by a related variable having the χ^2 distribution with one degree of freedom. (χ, then, is a unit normal variable.) Yates further points out that, for relatively small fourfold tables, the approximation is much improved by correction for continuity, i.e., for the replacement of the discrete hypergeometric variable by a continuous variable. χ^2 in the fourfold table above is given by

$$\chi^2 = \frac{N(mN - Mn)^2}{n(N-n)M(N-M)}. \tag{2}$$

If, in (2), the value of χ^2 at the 100α percentage point and the known values of the other quantities are inserted, the resulting expression is a quadratic equation in M. The discriminant of the quadratic form being positive except in the trivial case, $n=N$, the two real roots are easily interpreted. The smaller root, M_1, is the value for which $\chi(m)$ is at the upper 50α percentage point and, hence, m is approximately at the upper 50α percentage point of its (exact) hypergeometric distribution; a similar statement holds for M_2. The interval, $M_1 \leqq M \leqq M_2$, is then a confidence interval for M with confidence coefficient $(1-\alpha)$.

In what has been said immediately above, no allowance has been made for the continuity correction. Since, for the smaller root, m is in the upper tail of its distribution, the appropriate correction for continuity (following Yates) is subtraction of $1/2$; for the larger root, we use $m+\frac{1}{2}$. Thus, the confidence intervals for M, without and with correction for continuity, are, respectively,

$$\frac{N}{2k_\alpha}\left[k_\alpha + (2m-n) - \sqrt{k_\alpha^2 - \frac{2}{n}k_\alpha[m^2 + (n-m)^2] + (2m-n)^2}\right] \leqq M \leqq \frac{N}{2k_\alpha}\left[k_\alpha + (2m-n) + \sqrt{k_\alpha^2 - \frac{2}{n}k_\alpha[m^2 + (n-m)^2] + (2m-n)^2}\right] \tag{3}$$

(uncorrected)

and

$$\frac{N}{2k_\alpha}\Big[k_\alpha + (2m - n - 1) - \sqrt{k_\alpha^2 - \frac{2}{n}k_\alpha[(m-\tfrac{1}{2})^2 + (n-m+\tfrac{1}{2})^2] + (2m-n-1)^2}\Big] \tag{4}$$
$$\leqq M \leqq \frac{N}{2k_\alpha}\Big[k_\alpha + (2m - n + 1) + \sqrt{k_\alpha^2 - \frac{2}{n}k_\alpha[(m+\tfrac{1}{2})^2 + (n-m-\tfrac{1}{2})^2] + (2m-n+1)^2}\Big]$$

(corrected)

where $k_\alpha = n + (1 - n/N)\chi_\alpha^2$.

IV. SOME EXAMPLES

The three examples given below are chosen for their dissimilarity; they are not intended to be comprehensive.

(*i*) In a very small sample inquiry, we ask nine persons, randomly selected from a group of 100, whether they are in favor of a certain proposal and we find three in favor, six opposed. We wish to construct a 95 per cent confidence interval for the number, M, in the whole group, in favor of the proposal. Using (*a*) equation (3), (*b*) equation (4), and (*c*) exact hypergeometric methods, we obtain the following confidence intervals:

(*a*) $13 \leqq M \leqq 63$ (*uncorrected approximation*)

(*b*) $10 \leqq M \leqq 67$ (*corrected approximation*)

(*c*) $9 \leqq M \leqq 68$ (*exact*)

The interval (*b*) actually almost included the possibility of 68, since the cutoff value was 67.97. The Clopper and Pearson chart would indicate M between 7 and 72 inclusive.

(*ii*) This example is somewhat unrealistic and is chosen only to illustrate the use of existing tables and to point up the desirability of extending these tables. In this case, having observed that three out of eight randomly chosen from a group of 20 exhibit a characteristic, we seek a 90 per cent confidence interval for the number M, in the complete group of 20. Here, the χ^2 approximation, corrected for continuity, gives $3.023 < M < 13.042$. These values are in excellent agreement with the

exact computation based on the following values from the author's table [5]:

$$P_r\{m \geqq 3 \mid M = 3\} = .0491$$
$$P_r\{m \geqq 3 \mid M = 4\} = .1531$$
$$P_r\{m \leqq 3 \mid M = 13\} = .0521$$
$$P_r\{m \leqq 3 \mid M = 14\} = .0181$$

Examination of these probabilities indicates that, on the one hand, $M = 3$ is almost acceptable and, on the other, $M = 13$ is barely acceptable at the 90 per cent level.

(*iii*) Grant [4, pp. 412–14] considers in some detail a series of acceptance sampling plans in which 5 per cent samples are taken from lots of 2500 items. In this series of plans, the acceptance numbers are 0, 1, 2, 3 and 4, respectively. We now wish to construct 95 per cent confidence intervals for the number of defectives in the lot of 2500 when we observe 0, 1, 2, 3 or 4 defectives in the sample of 125. The table below gives confidence intervals obtained by the corrected χ^2 approximation.

95 per cent confidence intervals for number of defectives in lot of 2500 when sample size is 125.

Number of defectives in sample	Confidence interval (inclusive)
0	0– 89
1	2–122
3	8–152
3	16–180
4	27–207

Similar computations were carried out in this case for the normal approximation to the binomial approximation to the hypergeometric distribution. In every instance, a somewhat wider confidence interval was obtained; the results, however, were in excellent agreement in the sense that the intervals in the table above fell wholly within the second set.

In the case of the last example, it was possible to compare the confidence intervals obtained by our procedure with those given by the Chung and DeLury charts.[1] For each of the five acceptance numbers above, the Chung and DeLury confidence interval partially overlaps

the interval given above. For the acceptance number 3, for example, their confidence interval is (12, 170) as against our (16, 180). Exact computation of the lower confidence limit places it at 14, however, while the exact upper limit is at 169.

REFERENCES

[1] Clopper, C. J., and Pearson, Egon S., "The use of confidence or fiducial limits illustrated in the case of the binomial," *Biometrika*, 26 (1934), 404–13.

[2] Deming, William Edwards, *Some Theory of Sampling*, New York, John Wiley and Sons, 1950.

[3] Finney, D. J., "The Fisher-Yates test of significance in 2×2 contingency tables," *Biometrika*, 35 (1948), 145–56.

[4] Grant, Eugene L., *Statistical Quality Control*, New York, McGraw-Hill Book Co., 1946.

[5] Katz, Leo, *Tables of the hypergeometric cumulative distribution function:* unpublished, 1942.

[6] Yates, F., "Contingency tables involving small numbers and the χ^2 test," *Supplement to the Journal of the Royal Statistical Society*, 1 (1934), 217–35.

[7] Chung, J. H., and DeLury, D. B., *Confidence Limits for the Hypergeometric Distribution*, University of Toronto Press, 1950.

THE UP-AND-DOWN METHOD WITH SMALL SAMPLES*

K. A. Brownlee, J. L. Hodges, Jr., and Murray Rosenblatt
University of Chicago

I. INTRODUCTION

The up-and-down method for estimating the 50 per cent response point of quantal data was originally devised for testing the sensitivity of explosives. Dixon and Mood [6] suggested that this method might have advantages in other fields of application. They gave approximate maximum likelihood estimates for the parameters of the (normal) response curve, and approximate formulas for the standard errors of these estimates. They pointed out that, on the basis of formulas for the asymptotic variance, the up-and-down method was 30 or 40 per cent more efficient than the usual probit analysis method.

In spite of this efficiency advantage, the up-and-down method does not seem to have been given much consideration in such fields as bioassay [10, p. 909] or fatigue testing of metals. It has suffered from two restrictions which, taken together, have greatly limited its usefulness in many experimental situations:

(a) The trials must be made sequentially; before each trial is started, the response to the preceding trials must be known.
(b) The efficiency advantage has not been explored for small samples: Dixon and Mood express the fear (p. 112), that "measures of reliability may well be very misleading if the sample size is less than forty or fifty."

Thus, the minimum duration of the experiment would have to be forty or fifty times the mean time required for response. This disadvantage would, in many cases, more than outweigh the efficiency advantage.

The main content of this paper is a report on computations made to determine the actual performance of some estimates for the mean dosage parameter μ, based on up-and-down series of length 10 or less. It is found that the Dixon-Mood formula for the asymptotic variance is reasonably reliable even in samples as small as 5 to 10. Thus, restriction (b) is not necessary. As a consequence of this fact, the design of up-and-down experiments may be altered, so that several independent series are run simultaneously, without serious loss of accuracy. In this way, restriction (a) can be considerably reduced. Finally, the possibility of conducting an experiment with several short series run in

* This work was supported by the Office of Naval Research.

parallel introduces a flexibility into the design, which permits us to take advantage of special features of some experimental situations with a still further increase in efficiency.

We have not considered the problem of estimating the scale parameter σ. The reason for this is partly that μ is usually the parameter of greater interest, but primarily that with small samples no estimate for σ can be accurate enough to have much value. Even if μ were known, and even if the trials are conducted at stimuli giving the most efficient estimation [9], over 200 trials would be required to estimate σ within 20 per cent with confidence of 95 per cent. Our experience is that in most experimental situations, the scale parameter is sufficiently stable that the experimenter can guess its value in advance from past experience more accurately than he can estimate it from a small sample [2, p. 476]. Fortunately, our procedures require only that σ be known within rough limits, and the performance of the estimates for μ are not sensitive to errors in the guessed value of σ.

II. THE ESTIMATE $\hat{\mu}$

We assume that the stimulus scale has been so chosen that the response curve is an integrated normal curve with parameters μ and σ. That is, the probability of response P to stimulus y is given by

$$P = \int_{-\infty}^{(y-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

We fix a system of equally spaced stimuli, $y_0 \pm id$, $i = 0, 1, 2, \cdots$, and perform the first trial at stimulus y_0. Each subsequent trial is performed at stimulus d units below or above that of the immediately preceding trial, according as the immediately preceding trial did or did not evoke a positive response. We may without loss of generality choose our scale so that $\sigma = 1$, and we do this in order to simplify notation.

The series of stimuli used in an up-and-down experiment form a stochastic process, whose principal feature of interest is that the stimuli tend to have a distribution centered at μ. In a series of length n, a natural estimate for μ is simply the arithmetic mean of the n stimuli used. In such a series, however, the first stimulus y_0 does not contain any "information" about μ, since it was chosen in advance by the experimenter. On the other hand, the level of stimulus for the $(n+1)$st trial is informative and is known, even though this trial is not performed. Let C denote the sum of the stimuli used in trials 2, 3, $\cdots$, $n+1$; we shall term this the *score*. As our first estimate for μ we shall consider $\hat{\mu} = C/n$.

The estimate $\hat{\mu}$ is not quite the same as that given by Dixon and Mood, but the two are easily seen to be equivalent asymptotically, so we may use for $\hat{\mu}$ the asymptotic variance formula derived in [6]. This is $\sigma_A^2 = 2G^2/n$, where G is a function of d and $y_0 - \mu$. Dixon and Mood give a graph of the function $G(d, y_0 - \mu)$; a few representative values are tabled below.

d	$G(d, 0)$	$G(d, d/2)$
0	0.886	0.886
2/3	0.961	0.961
1	1.004	1.004
3/2	1.076	1.102

For given d, the extremes of G will occur approximately at $y_0 - \mu = 0$ and at $y_0 - \mu = d/2$. It appears from this table that for $d \leqq 3/2$, G does not sensibly depend on its second argument. This means that the asymptotic accuracy of the estimate does not depend much on the "phasing" of μ with the system of stimuli.

Dixon and Mood recommend that d be chosen approximately equal to σ. The experimenter must guess σ and then set d equal to his guessed value. The table of G shows that he may err considerably in his guess for σ without much affecting the asymptotic accuracy of the estimate for μ.

Our concern is with the actual accuracy of $\hat{\mu}$ for small n. We shall gauge this by means of the error variance $E(\hat{\mu} - \mu)^2$, which is the natural generalization of the variance when biased estimates are being considered. $E(\hat{\mu} - \mu)^2$ depends on n, $y_0 - \mu$, and d. It may be computed recursively. Let $C_n(y_0, \mu, d)$ denote the score obtained in an up-and-down series of length n, with step size d, started at y_0. Let P_0 denote the probability of positive response at stimulus y_0. We have

$$C_{n+1}(y_0, \mu, d) = \begin{cases} y_0 + d + C_n(y_0 + d, \mu, d) & \text{if the first trial fails} \\ y_0 - d + C_n(y_0 - d, \mu, d) & \text{if the first trial succeeds} \end{cases}$$

therefore

$$E[C_{n+1}(y_0, \mu, d)] = E[C_1(y_0, \mu, d)] + P_0 E[C_n(y_0 - d, \mu, d)] + (1 - P_0) E[C_n(y_0 + d, \mu, d)]$$

and

$$
\begin{aligned}
E[C_{n+1}(y_0, \mu, d) - (n+1)\mu]^2 &= E[y_0 + d - \mu + C_n(y_0 + d, \mu, d) - n\mu]^2(1 - P_0) \\
&\quad + E[y_0 - d - \mu + C_n(y_0 - d, \mu, d) - n\mu]^2 P_0 \\
&= E[C_1(y_0, \mu, d) - \mu]^2 + 2(y_0 - \mu)\{E[C_{n+1}(y_0, \mu, d) \\
&\quad - (n+1)\mu] - E[C_1(y_0, \mu, d) - \mu]\} \\
&\quad + 2d\{(1 - P_0)E[C_n(y_0 + d, \mu, d) - n\mu] \\
&\quad - P_0 E[C_n(y_0 - d, \mu, d) - n\mu]\} \\
&\quad + (1 - P_0)E[C_n(y_0 + d, \mu, d) - n\mu]^2 \\
&\quad + P_0 E[C_n(y_0 - d, \mu, d) - n\mu]^2.
\end{aligned}
$$

TABLE 1

$E(\hat{\mu}) - \mu$ AND $E(\hat{\mu} - \mu)^2$, $d = 1$

n \ $y_0 - \mu$	0	.5	1	1.5	2	2.5	3	3.5	4	σ_A^2
1	0	.117	.317	.634	1.046	1.512	2.003	2.500	3.000	2.016
	1.000	.867	.635	.651	1.182	2.312	4.016	6.253	9.000	
2	0	.116	.242	.422	.701	1.079	1.525	2.007	2.501	1.008
	.567	.284	.582	.557	.697	1.245	2.354	4.034	6.258	
3	0	.072	.180	.331	.529	.792	1.136	1.553	2.017	.672
	.477	.365	.431	.331	.559	.813	1.387	2.449	4.081	
4	0	.068	.142	.249	.413	.632	.899	1.219	1.602	.504
	.372	.320	.378	.325	.423	.536	.968	1.592	2.621	
5	0	.047	.116	.212	.337	.506	.732	1.006	1.319	.403
	.328	.284	.311	.284	.362	.428	.700	1.100	1.843	
6	0	.047	.097	.170	.282	.432	.616	.839	1.110	.336
	.256	.254	.279	.252	.300	.359	.553	.824	1.346	
7	0	.034	.084	.153	.243	.366	.529	.728	.956	.288
	.249	.227	.240	.228	.267	.302	.443	.657	1.042	
8	0	.035	.073	.128	.213	.326	.464	.633	.838	.252
	.218	.206	.220	.205	.231	.266	.375	.529	.827	
9	0	.026	.065	.137	.189	.285	.412	.568	.746	.224
	.200	.187	.195	.194	.211	.232	.317	.449	.683	
10	0	.033	.059	.103	.170	.277	.371	.507	.671	.202
	.180	.175	.178	.171	.189	.220	.281	.379	.570	

TABLE 2

$E(\hat{\mu})-\mu$ AND $E(\hat{\mu}-\mu)^2$, $d=2/3$

n \ $y_0-\mu$	0	2/3	4/3	2	8/3	σ_Δ^2
1	0	.337	.788	1.364	2.005	1.848
	.444	.449	.768	1.899	4.027	
2	0	.268	.609	1.090	1.687	.924
	.336	.377	.564	1.260	2.862	
3	0	.215	.491	.877	1.398	.617
	.296	.315	.449	.883	1.993	
4	0	.177	.403	.725	1.162	.462
	.259	.278	.364	.663	1.420	
5	0	.149	.339	.609	.984	.370
	.234	.244	.308	.515	1.060	
6	0	.128	.290	.522	.845	.308
	.209	.220	.264	.418	.816	
7	0	.111	.253	.455	.737	.264
	.191	.197	.233	.347	.652	
8	0	.098	.223	.401	.651	.231
	.174	.180	.206	.297	.533	
9	0	.087	.199	.358	.581	.205
	.161	.164	.186	.258	.447	
10	0	.079	.180	.323	.525	.185
	.148	.152	.169	.228	.381	

These formulas enable us to compute $[E(\hat{\mu})-\mu]$ and $E(\hat{\mu}-\mu)^2$. Beginning with $n=1$ we have

$$E(\hat{\mu}) - \mu = E[C_1(y_0, \mu, d)] - \mu = y_0 - \mu + d(1 - 2P_0)$$

and

$$\begin{aligned} E(\hat{\mu} - \mu)^2 &= E[C_1(y_0, \mu, d) - \mu]^2 \\ &= (y_0 - \mu + d)^2(1 - P_0) + (y_0 - \mu - d)^2 P_0. \end{aligned}$$

TABLE 3

$E(\hat{\mu})-\mu$ AND $E(\hat{\mu}-\mu)^2$, $d=3/2$

n \ $y_0-\mu$	0	3/2	3	9/2	6	σ_A^2
1	0 2.225	.200 .601	1.504 2.274	3.000 9.000	4.500 20.240	2.317
2	0 .863	.150 .864	.854 .881	2.252 5.075	3.750 14.063	1.158
3	0 .734	.105 .527	.604 .682	1.569 2.530	3.001 9.010	.772
4	0 .496	.080 .494	.456 .503	1.203 1.625	2.302 5.337	.579
5	0 .445	.064 .371	.366 .449	.965 1.120	1.862 3.582	.463
6	0 .347	.054 .346	.305 .350	.805 .867	1.554 2.546	.386
7	0 .320	.046 .282	.262 .322	.690 .665	1.333 1.938	.331
8	0 .267	.040 .266	.229 .268	.604 .559	1.167 1.506	.290
9	0 .250	.036 .227	.204 .251	.537 .459	1.037 1.229	.257
10	0 .216	.032 .216	.183 .217	.483 .404	.933 1.009	.232

From these the moments for $n>1$ follow recursively by using the above formulae.

Tables 1, 2, and 3 give $E(\hat{\mu})-\mu$ and $E(\hat{\mu}-\mu)^2$ for $d=1$, 2/3, and 3/2 respectively, and for $n=1(1)10$. In Table 1, $y_0-\mu=0(1/2)$ 3, while in Tables 2 and 3, $y_0-\mu=0(d)$ $3d$. For comparison we give also the asymptotic variance of $\sigma_A^2=2G^2(d,\ d/4)/n$. We see that in all tables, if $y_0-\hat{\mu}\leqq 2d$, the actual $E(\hat{\mu}-\mu)^2$ is less than σ_A^2, but that as $y_0-\mu$ increases beyond about $(5/2)d$, $E(\hat{\mu}-\mu)^2$ rapidly increases, particularly for small d. This means that the asymptotic variance gives reliable or

even conservative estimates for the accuracy of $\hat{\mu}$ in small samples, provided the experimenter can manage to start the process within, say, two testing intervals, $2d$, of the mean, μ.

III. THE ESTIMATE μ'

In many experimental situations the experimenter will be able to guess the value of μ to within 2 steps, where the step is the guessed value of σ. In these situations the estimate $\tilde{\mu}$ is quite satisfactory, being both simple and accurate. But it is always possible that the guess will be badly in error, and we now present a modified up-and-down procedure which, by sacrificing some efficiency when $y_0-\mu$ is small, provides an estimate with a guaranteed accuracy independent of $y_0-\mu$. It is clear that if we start an up-and-down series at too high a stimulus, we shall very probably obtain only positive responses until the stimulus level has been reduced to a value near μ. Similarly, if y_0 is far below μ, this fact will probably reveal itself, through an initial run of negative responses. This suggests that we should interpret an initial run of responses of the same sign as an indication that y_0 was badly chosen, and that these trials may well be ignored in estimating μ.[1]

We are thus led to formulate the following design. Choose n in advance, and continue the experiment until there have been $n-1$ trials in addition to the trials in the initial run of constant sign. Thus, the total number N of trials is a random variable, $N \geqq n$, with $N=n$ only if the first and second trials give contrary responses. As score, we use C', the sum of the stimuli on the last $n-1$ trials and on the $(N+1)$st trial (which is, of course, not performed). For our estimate, we use $\mu'=C'/n$.

To illustrate, let $d=1$, $y_0=0$, $n=3$. Suppose that the first two trials failed, but the third and fourth trials succeeded. This means that the successive stimulus levels were 0, 1, 2, and 1, and if a fifth trial was to be performed, it would be at level 0. Here $N=4$, $C'=2+1+0=3$, $\mu'=3/3=1$.

Values of $E(\mu')-\mu$ and $E(\mu'-\mu)^2$ can be computed from those for $E(\tilde{\mu})-\mu$ and $E(\tilde{\mu}-\mu)^2$. We first compute the probability of each possible initial run. Each such run leads to a specific starting value for a series whose moments can be obtained from tables of $E(\tilde{\mu})-\mu$ and $E(\tilde{\mu}-\mu)^2$. We give in Table 4 values of $E(\mu')-\mu$ and $E(\mu'-\mu)^2$ for $d=2/3$, 1, 3/2, $y_0=0(d)\ 3d$, and $n=5$ and 10. In Table 5 we give the expected number of wasted observations, $E(N)-n$, for $d=2/3$, 1, 3/2 and $y_0-\mu=0(d)\ 3d$.

[1] The estimate recommended by Dixon and Mood also attempts to avoid using the results of initial trials made far from μ, but in a way different from that proposed here.

An appropriate measure for the comparison of two methods of estimation is provided by the inverse of the sampling variance V of the statistic being estimated. We may thus, following Fisher [8], define a quantity I such that $I=1/V$. In the case of the comparison of experiments involving different numbers of observations it is fairest to make the comparison on a per unit observation basis, so we use $I'=I/n=1/nV$.

In choosing between the estimates $\hat{\mu}$ and μ', several considerations should be kept in mind. If the initial guess for μ is fairly accurate, $\hat{\mu}$ will have the smaller error variance. In addition, μ' involves "wasted" observations. Comparing $nE(\hat{\mu}-\mu)^2$ with $E(N)E(\mu'-\mu)^2$ for $n=10$, $\hat{\mu}$ is seen to be more efficient than μ' when $y_0-\mu=0$, d, or $2d$ for $d=2/3$, 1, and 3/2. On the other hand, u' is more efficient for large values of $|y_0-\mu|$. Furthermore, in using μ' we have a known upper bound on the error variance, depending on d but independent of $y_0-\mu$. Also, the bias of μ' is bounded and is considerably smaller than that of $\hat{\mu}$ for $y_0-\mu=2d$, and $3d$.

An examination of Tables 4 and 5 suggests a further modification of the sampling method which may have desirable features in some experimental situations. We observe (a) that the smallest value of d gives

TABLE 4

$E(\mu'-\mu)$ AND $E(\mu'-\mu)^2$, $n=5$ AND 10

d	n	$y_0-\mu$				σ_A^2
		0	d	$2d$	$3d$	
2/3	5	0 .267	.158 .317	.267 .423	.317 .513	
	10	0 .157	.083 .167	.141 .199	.168 .224	.185
1	5	0 .357	.128 .410	.187 .485	.198 .516	
	10	0 .191	.065 .202	.095 .222	.101 .230	.202
3/2	5	0 .463	.074 .507	.096 .547	.097 .551	
	10	0 .238	.037 .240	.048 .251	.049 .252	.232

TABLE 5

EXPECTED NUMBER OF OBSERVATIONS USED IN APPROACHING μ BUT NOT USED IN THE ESTIMATE OF μ, $E(N)-n$

$y_0-\mu$	$d=2/3$	$d=1$	$d=3/2$
0	0.276	0.162	0.067
d	0.500	0.493	0.498
$2d$	1.115	1.300	1.429
$3d$	1.976	2.274	2.428

the smallest values for $E(\mu'-\mu)^2$ while (b) the largest value of d gives the smallest value of $E(N)$, for a given value of $y_0-\mu$. This is intuitively to be expected: a few big steps will get the process into the neighborhood of μ; but once there, small steps will keep the process closer to μ.

Now suppose we begin the process with big steps, but change to small steps with the first change of sign. This design may be expected to have the small $E(N)$ characteristic of big steps, and also the small $E(\mu'-\mu)^2$ characteristic of small steps. As an illustration, suppose we begin at $y_0-\mu=4.5$ with $d=3/2$, and at the first change of sign change to $d=2/3$. We shall have $E(N)=12.43$, and a simple computation (using graphical interpolation in Table 5) gives $E(\mu'-\mu)^2=0.22$, approximately. Thus $E(N)\ E(\mu'-\mu)^2=2.73$, approximately. Had we used $d=3/2$ throughout, we should have had $E(N)\ E(\mu'-\mu)^2=12.43\times0.252=3.13$, while the use of $d=2/3$ throughout would give $E(N)\ E(\mu'-\mu)^2=16.7\times0.224=3.74$, approximately. The change in step size has given a substantial reduction in $E(N)\ E(\mu'-\mu)^2$.

An elaboration of this idea would involve a series of step sizes, decreasing to zero. Such a model would have an asymptotic efficiency with $E(N)\ E(\mu'-\mu)^2=1.57$. In this connection, see [12].

IV. COMPARISON OF EFFICIENCIES OF UP-AND-DOWN METHOD WITH THAT OF PROBIT ANALYSIS

A comparison of the efficiencies of the estimates $\tilde{\mu}$ and μ' with that of the standard probit analysis is of interest but rather difficult. In the case of $\tilde{\mu}$ and μ' the accuracies depend on the choice of d, and for the probit method the accuracy of the estimate is markedly affected by the choice of the number of stimuli, their spacing, and the distribution of the trials among them. Also, the true error variance of the probit estimate does not seem to be easily obtainable.

We can, however, make the comparison on the following basis. For

the probit method, we shall use a very common design, namely five stimulus levels equally spaced at intervals d with equal numbers of trials at each level. We shall assume that the stimulus interval d represents the experimenter's guess for σ, and shall use this same value for the step-size in the up-and-down series.

We shall gauge the accuracy of the probit estimate by the usual approximate formula for the asymptotic variance σ_p^2 (equation 3.6 of [7]):

$$\sigma_p^2 = \frac{1}{m}\left[\frac{1}{\sum w_i} + \frac{\bar{x}^2}{\sum w_i x_i^2 - \bar{x}^2 \sum w_i}\right]$$

where x_i are the stimuli, w_i their probit weights, m the number of trials at each stimulus, and $\bar{x} = \sum w_i x_i / \sum w_i$.

In Table 6, column 3 gives $1/I'$ for the probit design specified above for $y_0 - \mu = 0$, d, $2d$, and $3d$ and $d = 2/3$, 1, and 3/2. From the formula given for σ_p^2 we calculated $m\sigma_p^2$ and then $n\sigma_p^2$, where $n = 5m =$ total number of observations. Columns 4 and 8 are obtained by abstracting the appropriate values of $E(\tilde{\mu} - \mu)^2$ from Tables 1, 2, and 3 and multiplying by $n = 5$ and 10. For columns 6 and 10, $E(N)$ is obtained from Table 5 and $E(\mu' - \mu)^2$ from Table 4. Column 5 gives R_1 defined as

$$R_1 = \frac{(I' \text{ for up-and-down estimate } \tilde{\mu})}{(I' \text{ for probit method})}$$

and column 11 gives R_2 defined similarly but for μ'. Thus nR_1 is the number of observations required with the probit method to give the same error variance as would be obtained from $\tilde{\mu}$ with n observations, and $E(N)R_2$ is the number of observations required with the probit method to give the same error variance as would be obtained from μ' with a number of observations whose expected value is $E(N)$.

All values of R_1 and R_2 exceed 1, indicating that in all the circumstances considered the probit method requires more observations. When $d = 1$ (which is the usual objective for d), the probit method requires at least 45 per cent more observations than μ', and at least 55 per cent more observations than $\tilde{\mu}$, for all values of $y_0 - \mu$ considered.

It might be felt that the above comparison is unfair to the probit method, in that we have judged its accuracy by an asymptotic formula. The only evidence on this point available to us, however, tends to show that the contrary is true. Berkson [2] has reported on an extensive sampling experiment one of whose purposes was to evaluate the small-sample performance of the maximum likelihood estimate for the parameters of the logistic function. This function is scarcely distinguish-

able from the integrated normal. For thirty trials evenly divided between three equally spaced levels, in estimating the location parameter with the scale parameter fixed, he found that the actual error variance of the maximum likelihood estimate was some 10 per cent greater than predicted by the asymptotic formula. (It should, however, be remarked that Berkson was able to attain or even slightly better the asymptotic variance with a minimum chi-square estimate.) While we realize that Berkson's results are not exactly comparable, they lead us to feel that the efficiency advantage shown for the up-and-down method in Table 6 is conservative. Further information may be forthcoming with the publication of [3].

TABLE 6

COMPARISON OF EFFICIENCIES OF $\hat{\mu}$ AND μ' WITH THAT OF A PROBIT ANALYSIS

d	$y_0-\mu$	$n\sigma_p^2$	$n=5$				$n=10$			
			$nE(\hat{\mu}-\mu)^2$	R_1	$E(N)E(\mu'-\mu)^2$	R_2	$nE(\hat{\mu}-\mu)^2$	R_1	$E(N)E(\mu'-\mu)^2$	R_2
2/3	0	2.11	1.17	1.81	1.41	1.50	1.48	1.43	1.61	1.31
	d	2.68	1.22	2.19	1.74	1.54	1.52	1.76	1.75	1.53
	2d	5.97	1.54	3.88	2.59	2.31	1.69	3.53	2.21	2.70
	3d	23.6	2.58	9.17	3.58	6.60	2.28	10.4	2.68	8.80
1	0	2.82	1.64	1.72	1.84	1.53	1.80	1.56	1.94	1.45
	d	3.13	1.56	2.01	2.25	1.39	1.78	1.76	2.12	1.48
	2d	6.98	1.81	3.86	3.06	2.29	1.89	3.70	2.51	2.78
	3d	63.0	3.50	18.0	3.75	16.8	2.81	22.4	2.82	22.3
3/2	0	4.15	2.22	1.87	2.62	1.58	2.16	1.92	2.54	1.64
	d	4.21	1.86	2.27	2.79	1.51	2.16	1.95	2.52	1.67
	2d	7.71	2.24	3.43	3.52	2.19	2.17	3.55	2.87	2.69
	3d	414	5.60	74.0	4.09	101	4.04	103	3.13	132

We remark that the sense of the comparison of Table 6 is not dependent on the particular probit design we chose. A probit design which concentrated the trials more tightly would gain in efficiency somewhat for y_0 near μ, but would pay for this by blowing up more quickly as the initial guess became bad; and conversely, a looser distribution of the trials would give fair performance ever a wide range of y_0 values, but would be even less efficient at y_0 near μ.

V. DESIGNS WHICH REDUCE THE LENGTH OF THE EXPERIMENT

The necessity of performing the trials sequentially is the principal objection to the use of the up-and-down method in many quantal response experiments. The total time required is likely to be prohibitive

unless the response to the stimulus is immediate, as for example in explosives sensitivity testing.[2]

We shall now investigate the possibility of mitigating this disadvantage by running several short up-and-down series simultaneously, rather than a single long series. Let there be k series, each of length n, with $m=kn$ the total number of trials. We shall conduct the series independently of each other, and use the same initial stimulus y_0 and the same step size d for all k series. From the ith series we form the estimate μ_i' as in section 3, and for our over-all estimate use

$$\bar{\mu} = \left(\sum_{i=1}^{k} \mu_i' \right) \Big/ k.$$

Clearly $\sigma_{\bar{\mu}}^2 = \sigma_{\mu'}^2/k$, while $E(\bar{\mu}) = E(\mu')$.
Thus

$$\begin{aligned} E(\bar{\mu} - \mu)^2 &= \sigma_{\bar{\mu}}^2 + [E(\bar{\mu}) - \mu]^2 \\ &= E(\mu' - \mu)^2/k + [E(\mu') - \mu]^2(k-1)/k. \end{aligned}$$

The error variance of $\bar{\mu}$ may be obtained at once from Table 4 for $d=2/3$, 1, 3/2, $n=5$ and 10.

To illustrate, consider four independent series each of length ten ($k=4$, $n=10$, $m=40$). The following table gives the error variance corresponding to several value of $y_0-\mu$, where we have taken $d=1$. We also give the values of $R_2=(40\sigma_p^2)/[4E(N)E(\bar{\mu}-\mu)^2]$.

$y_0-\mu$	$E(\bar{\mu}-\mu)^2$	R_2
0	0.048	1.45
1	0.054	1.39
2	0.062	2.48
3	0.065	19.7
4	0.066	—

[2] It should not, however, be thought that the total time required to run an up-and-down series is n times as great as the time required to conduct a probit experiment with n trials. In the first place, the greater efficiency of the up-and-down method means that fewer than n trials are needed with it: as is explained in Section 4, something like $2n/3$ trials will give the same accuracy as is obtained from a probit analysis based on n trials. Secondly, in many experiments, particularly in biological ones, the response time is not fixed but variable. The time required for the probit experiment is thus the *maximum* of n response times while that required for the up-and-down method is the sum of $2n/3$ response times. To illustrate, if response times have a chi-square distribution with six degrees of freedom (this distribution fits fairly well a series of response times published by Perry [11, p. 40]) and if $n=30$, the median time required to complete the up-and-down experiment will be 7.6 times as long as that for the probit experiment, rather than 30 times as long as one might at first think.

For a single series of length 40, the variance as given by the asymptotic formula is $1.008/20 = 0.050$. It appears that there is no appreciable loss of efficiency in breaking a series of 40 into four series of 10 provided that a good guess for μ is made. We are also assured of a low bound on the error variance even if the guess is bad. For all values of $y_0 - \mu$ considered, the ratio R_2 is at least 1.30. This indicates that a 5-level probit design would require at least 30 per cent more observations to yield an estimate of the accuracy of $\bar{\mu}$, even for the most favorable choice of y_0.

VI. COMPARATIVE EXPERIMENTS

In much work of the type considered in this paper, we do not make an absolute determination but rather a comparison of an unknown with some standard. What is desired in these situations is an estimate for the difference between the values of μ associated with two treatments. For example, in bioassay we may wish to estimate the amount by which the LD 50, μ_1, of a new drug falls below that, μ_2, of a standard drug. It is clear that the up-and-down method may be adapted to this kind of problem. We can simply run two series, one for each drug, and take $\hat{\mu}_1 - \hat{\mu}_2$ as our estimate for $\mu_1 - \mu_2$. The possibility of running several short up-and-down series opens up additional ways of economizing on sample sizes.

It is well known that the response of an animal to a given dose is affected by (amongst other things) its weight, age, and previous history. The usual method of avoiding these disturbances is to select animals as homogeneous as possible, or, in the case of weight, to make the log dose simply proportional to the log weight. While in some cases this procedure will achieve a satisfactory compensation for the varying animal weights there are others where the relationship is either unknown or less simple. It is not safe to assume simple proportionality with a regression coefficient of unity, as regression coefficients of log dose on log weight as diverse as 0.573 [5] and 1.511 [4] have been reported.

In an orthodox probit comparative experiment, stratification with respect to some concomitant variable, such as animal weight, is possible, but it is not practical to take account of it in the analysis. In an up-and-down comparative experiment, however, the whole series of say 80 animals can be broken up on the basis of weight into say 4 groups of 20, and 10 animals from each such group allocated at random to the standard and 10 to the unknown. This would give us more homogeneous material and a correspondingly reduced standard error.

Due to a paucity of suitable data in the literature, it is difficult to predict what magnitude of gain may be expected through this sub-

division of the material. However, as an example we may take Behren's data [1] on digitalis in frogs. For February 1928 he measured the fatal dose of digitalis for 38 frogs varying in weight from 21.8 to 37.5 grams, and as usual expressed the dose as a fraction of the frogs' weight. However, a plot of dose, so defined, against frogs' weight discloses that the heavier frogs require a smaller dose. If we rank the frogs in order of increasing weight and break up the group of 38 into 4 groups of 8, 10, 10, and 10, then the variance of dose in his units within these groups is 0.0035550 compared with an over-all variance of 0.0046134. The ratio of these variances is 0.771: thus 7.7 animals when used in the smaller groups will give the same information as 10 animals in the complete group.

In the case of relatively large experiments, of course, the possibility exists of obtaining still greater homogeneity by classifying the animals into groups on the basis of more than one parameter.

A further possibility is the re-use of survivors from previous experiments. In general, we would expect the tolerance to change, and the inclusion of these animals in a comparative probit experiment would increase the heterogeneity of the material. However, in an up-and-down comparative experiment broken up into small groups, the survivors from earlier experiments could form some of the groups. This procedure if used with discretion should allow a saving of up to 50 per cent in the usage of animals with the up-and-down method. There is also, of course, a further reduction compared with probit experiments due to the relatively greater efficiency of the up-and-down method.

VII. CONCLUDING REMARKS

In choosing between the standard probit method and the up-and-down method for estimating μ, the following points may be kept in mind.

(1) The probit method will require substantially more observations. It shows up best, relatively, when the guess for μ is accurate, but even here (using a standard 5-level probit design with stimulus interval equal to one standard deviation) the probit method requires at least 50 per cent more trials than does the up-and-down estimate $\hat{\mu}$ in series of length 10. The comparison becomes rapidly less favorable to the probit method if the initial guess is not accurate. When the central stimulus is off-center by two standard deviations, the probit method requires approximately 3 times as many observations as ...

(2) With the up-and-down method applied sequentially, it is possible to obtain an estimate μ' with a guaranteed accuracy, regardless of

initial guess for μ (but depending, as does the probit method, on the guess for σ). This estimate, while less efficient than $\bar{\mu}$ for small $y_0-\mu$, is still more efficient than the probit method, which requires at least 40 per cent more trials on the average to give the same accuracy, when the guess for σ is correct.

(3) The small-sample performance of the up-and-down estimates is known, whereas that of the probit method is unknown and difficult to compute. The above comparisons are made on the basis of the asymptotic variance formula for the probit method, which is probably optimistic.

(4) The up-and-down estimates are easy to compute arithmetically. The probit estimate must be computed either by a troublesome iterative method, or by a graphical method involving the judgement (and hence the bias) of the computer.

(5) The up-and-down estimates always exist, whereas the probit estimate may not exist. In small samples with $y_0-\mu$ large, there is a substantial probability that the probit estimate will not exist.

(6) The probit method has the advantage that all trials may be performed simultaneously, whereas the up-and-down method requires that trials be made sequentially. We may, however, run the up-and-down trials in a number of parallel short series, without much loss of accuracy. In a typical situation, the up-and-down experiment may take 2 to 5 times as long to complete as the probit experiment.

(7) The up-and-down method, when arranged in short series in an experiment to compare treatments, can take advantage of any classification of experimental material designed to reduce the variance. This may lead to a further substantial increase in efficiency relative to the probit method.

In summary, we believe that the up-and-down method will prove superior to the probit method in any situation wherein the arrangement of the trials in series is not prohibited by experimental or cost considerations. Certainly there are many laboratories now using the probit method which would profit from a change to the up-and-down design.

Our thanks are due to Mr. J. McKibben for much of the computation included in this paper.

REFERENCES

[1] Behrens, B., "Zur Auswertung der Digitalisblatter in Froschversuch," *Archiv. fur Experimentelle Pathologie und Pharmakologie*, 140 (1929), 237–56.

[2] Berkson, Joseph, "Relative precision of minimum chi-square and maximum likelihood estimates of regression coefficients," in Neyman, Jerzy (editor),

Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1951, 471–79.

[3] Berkson, Joseph, "Relative precision of least squares and maximum likelihood estimates of regression coefficients," Abstract, *Annals of Mathematical Statistics*, 23 (1952), 148.

[4] Bliss, C. I., "The size factor in the action of arsenic upon silkworm larvae," *Journal of Experimental Biology*, 13 (1936), 95–110.

[5] Dawson, W. T., "Relation between age and weight and dosage of drugs," *Annals of Internal Medicine*, 13 (1940), 1594–615.

[6] Dixon, W. J., and Mood, A. M., "A method for obtaining and analyzing sensitivity data," *Journal of the American Statistical Association*, 43 (1948), 109–26.

[7] Finney, D. J., *Probit Analysis*, Cambridge, England, Cambridge University Press, 1947.

[8] Fisher, R. A., "The logic of inductive inference," *Journal of the Royal Statistical Society*, 98 (1935), 39–54.

[9] Friedman, Milton, "Planning an experiment for estimating the mean and standard deviation of a normal distribution from observations on the cumulative distribution," Chapter 11 in Statistical Research Group, Columbia University, *Techniques of Statistical Analysis*, New York, McGraw-Hill, 1947.

[10] Miller, Lloyd C., "Biological assays involving quantal responses," *Annals of the New York Academy of Sciences*, 52 (1950), art. 6, 903–19.

[11] Perry, W. L. M., "The design of toxicity tests," *Medical Research Council Special Report Series*, No. 270, London (1950).

[12] Robbins, Herbert, and Monro, Sutton, "A stochastic approximation method," *Annals of Mathematical Statistics*, 22 (1951), 400–07.

DESIGNING SINGLE-SAMPLING INSPECTION PLANS WHEN THE SAMPLE SIZE IS FIXED

ABRAHAM GOLUB*

Ballistic Research Laboratories, Aberdeen Proving Ground

A simple technique is developed for determining "best" single sampling inspection plans when the sample size is fixed. Two cases are considered: (1) single sampling plans for placing a lot into one of two categories and (2) single sampling plans for placing a lot into one of three categories. For case (1), the "best" plans are based upon a criterion of minimizing the sums of the producer's and consumer's risks. A similar criterion is employed as a basis for choosing "best" plans for case (2). The more general result for m categories is given also. Tables 1 through 8 are presented to enable the user to choose the appropriate sampling plans.

INTRODUCTION AND SUMMARY

THERE have been various approaches to the problem of designing single-sampling plans for acceptance sampling by attributes. It is generally accepted, however, that the primary objectives are twofold: (1) to accept at least a stated high proportion, say $1-\alpha$, of lots having true fraction defective not exceeding a given acceptable quality, say p_1 and (2) to accept not more than a stated small proportion, say β, of lots having true fraction defective not less than a given objectionable quality, say p_2. The quantities α and β are commonly known as the producer's and consumer's risks, respectively. The problem of selecting a single sampling plan is one of determining the parameter n (sample size) and the parameter c (acceptance number) which will simultaneously attain both of the aforementioned objectives. Grubbs has solved this problem and by using his tables it is quite simple to choose the single sampling plan which gives the desired protection.[1]

There are many occasions in practice where for economic, administrative, or practical reasons n is fixed and small (less than 50). For these cases, the attainment of the desired protection is usually impossible. A plan could be devised in which α or β is controlled; however, this may unnecessarily penalize either the consumer or producer. With n fixed, therefore, the problem is one of choosing a "best" compromise plan based upon some criterion. This paper proposes a solution of this problem.

* The author wishes to express his appreciation to Dr. F. E. Grubbs for reviewing this paper and for his continued interest and stimulating comments throughout its preparation.

[1] Grubbs was not necessarily the first to consider this problem; his tables, however, give his solution important advantages. See Frank E. Grubbs, "On designing single-sampling inspection plans," *Annals of Mathematical Statistics*, 20 (1949), 242–56.

The criterion which is employed as a basis for a "best" plan is that the sum of the probabilities of accepting lots of true quality p_1 (acceptable quality) and rejecting lots of true quality p_2 (objectionable quality) be a maximum. Mathematically, the expression which is to be maximized may be written as

$$P = Pr_{p_1}(A) + Pr_{p_2}(R) \tag{1}$$

where $Pr_{p_1}(A)$ represents the probability of accepting lots of true quality $= p_1$ and $Pr_{p_2}(R)$ represents the probability of rejecting lots of true quality $= p_2$. It is noted that in employing this criterion α and β may vary, whereas in the basic problem α and β are set values. For fixed n, the value of c which maximizes (1) is the integer nearest to

$$c = -\frac{1}{2} + \frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1}, \tag{2}$$

where $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$.

This paper also considers the problem of obtaining single sampling plans involving two acceptance parameters, c_1 and c_2, for the purpose of placing a lot into one of three categories, say 1, 2, or 3. Three quality levels are set; p_1 for category 1, p_2 for category 2, and p_3 for category 3. In this case a similar criterion for a "best" plan is employed. Mathematically the expression which is to be maximized may be written as

$$P = Pr_{p_1}(1) + Pr_{p_2}(2) + Pr_{p_3}(3) \tag{3}$$

where $Pr_{p_1}(1)$ represents the probability of placing lots of quality p_1 into category 1, $Pr_{p_2}(2)$ represents the probability of placing lots of quality p_2 into category 2, and $Pr_{p_3}(3)$ represents the probability of placing lots of quality p_3 into category 3. This criterion obviously minimizes the sum of the probabilities of misclassification.

The solutions for the c_1 and c_2 which maximize (3) are the integers nearest to

$$c_1 = -\frac{1}{2} + \frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1} \tag{4}$$

$$c_2 = -\frac{1}{2} + \frac{n}{\log \frac{p_3}{p_2} \Big/ \log \frac{q_2}{q_3} + 1}. \tag{5}$$

For the general case of classification into m categories, the solution for the c_i are

$$c_i = -\frac{1}{2} + \frac{n}{\log \frac{p_{i+1}}{p_i} \Big/ \log \frac{q_i}{q_{i+1}} + 1}. \tag{6}$$

The tables at the end of this paper are sampling tables which enable the user to select the proper acceptance numbers for single sampling, based upon the solution developed herein.[2] The sample size, n, ranges from five through forty in intervals of five. The lot qualities range from .01 to .20 for acceptable qualities and up to .40 for objectionable qualities. Lot qualities are tabulated in intervals of .01.

Examination of formulas (2), (4) and (5) indicates that these same tables can be used also to choose c_1 and c_2 in obtaining solutions for classification into three categories. This is accomplished by repeated entry into the tables, first with p_1 and p_2 to obtain c_1 and then with p_2 and p_3 to obtain c_2. In a similar manner, the c_i for the general case of m categories can be obtained from the tables.

The solutions developed in this paper are based upon the assumption of a Binomial Population.

In summary, tables are provided which enable the user to determine quickly the acceptance numbers, c, for fixed n, so that the sum of the probabilities of misgrading or misclassifying is a minimum. For those cases where n exceeds the range of the tables, a simple computation employing formula (2) yields the desired results.

---

[2] These tables originally appeared in this author's "Acceptance Numbers for Placing a Lot From Which a Single Sample is Drawn into One of Three Grades," Memorandum Report No. 581, Ballistic Research Laboratories, Aberdeen Proving Grounds, Md. (Nov. 1951).

EXAMPLES

I. Given:

$$n = 40$$
$$p_1 = .01$$
$$p_2 = .08.$$

Find c so that $\alpha+\beta$ is a minimum.

The last table yields $c=1$ when entered with $p_1=.01$ and $p_2=.08$. Thus the single sampling plan which minimizes $\alpha+\beta$ is $n=40$; $c=1$. Specifically, $\alpha=.06$; $\beta=.16$ and $\alpha+\beta=.22$.

II. Given:

$$n = 25$$
$$p_1 = .01$$
$$p_2 = .10$$
$$p_3 = .30.$$

Find c_1 and c_2 so that $Pr_{p_1}(1)+Pr_{p_2}(2)+Pr_{p_3}(3)$ is a maximum.

The fifth table yields $c_1=0$ when entered with $p_1=.01$ and $p_2=.10$ and also $c_2=4$ when entered with $p_2=.10$ and $p_3=.30$. Specifically, $Pr_{p1}(1)=.778$; $Pr_{p2}(2)=.830$; and $Pr_{p3}(3)=.910$ and their sum is 2.518.

THEORETICAL BASIS FOR CONSTRUCTION OF TABLES 1 THROUGH 8

Assuming an acceptable quality level of p_1 and an objectionable quality level of $p_2(p_1<p_2)$ and the sample size n fixed, we can write

$$Pr_{p_1}(A) = \sum_{i=0}^{c} \frac{n!}{i!(n-i)!} p_1^i q_1^{n-i} \tag{7}$$

$$Pr_{p_2}(R) = 1 - \sum_{i=0}^{c} \frac{n!}{i!(n-i)!} p_2^i q_2^{n-i} \tag{8}$$

where, assuming a binomial population, $Pr_{p_1}(A)$ represents the probability of accepting lots of true quality p_1 and $Pr_{p_2}(R)$ represents the probability of rejecting lots of true quality p_2. The criterion for a "best" plan, as outlined in the introduction, is that

$$P = Pr_{p_1}(A) + Pr_{p_2}(R) = \text{maximum}. \tag{9}$$

For c to be the acceptance number that maximizes (9), provided that

(9) has but one maximum, the following functional relationships must hold:

$$P_c - P_{c-1} > 0 \tag{10}$$

$$P_{c+1} - P_c < 0 \tag{11}$$

where the subscript denotes the acceptance number. Expanding (10) and (11) yields

$$\frac{n!}{c!(n-c)!}(p_1^c q_1^{n-c} - p_2^c q_2^{n-c}) > 0 \tag{12}$$

$$\frac{n!}{(c+1)!(n-c-1)!}(p_1^{c+1} q_1^{n-(c+1)} - p_2^{c+1} q_2^{n-(c+1)}) < 0. \tag{13}$$

Dividing both sides of the inequality (12) by the positive quantity $n!/[c!(n-c)!]$ gives

$$p_1^c q_1^{n-c} - p_2^c q_2^{n-c} > 0 \tag{14}$$

which is equivalent to

$$\left(\frac{p_1 q_2}{q_1 p_2}\right)^c q_1^n > q_2^n. \tag{15}$$

Since (p_1q_2/q_1p_2) is less than unity, taking logarithms of both sides of (15) shows that c is an integer satisfying

$$c < \frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1}. \tag{16}$$

Similarly (13) can be reduced to show that c is an integer satisfying

$$c > \frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1} - 1. \tag{17}$$

Combining (16) and (17), we find that c is an integer which satisfies

$$\frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1} - 1 < c < \frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1}. \tag{18}$$

From (18) it is obvious that c may be obtained by evaluating

$$-\frac{1}{2} + \frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1} \tag{19}$$

and rounding to the nearest integer. If the value of (19) falls exactly midway between two integers then equal maxima exist at those two integers.

An alternate and interesting approach for obtaining the acceptance number c which maximizes (9) is to approximate (9) as follows:

$$\begin{aligned} P \cong\ & \frac{n!}{0!n!} p_1^0 q_1^n + \frac{n!}{1!(n-1)!} p_1^1 q_1^{n-1} + \cdots \\ & + \int_{j-1/2}^{c+1/2} \frac{n!}{x!(n-x)!} p_1^x q_1^{n-x} dx \\ & + 1 - \left(\frac{n!}{0!n!} p_2^0 q_2^n + \frac{n!}{1!(n-1)!} p_2^1 q_2^{n-1} + \cdots \right. \\ & \left. + \int_{j-1/2}^{c+1/2} \frac{n!}{x!(n-x)!} p_2^x q_2^{n-x} dx \right) \end{aligned} \tag{20}$$

where j may be set arbitrarily equal to 0, or 1, or 2, $\cdots$, but at some value less than c. The range of integration for a particular term, say the $(k+1)$st goes from $k-\frac{1}{2}$ to $k+\frac{1}{2}$, thus employing a useful device for passing from a discrete variable to a continuous variable. Differentiating (20) with respect to c and equating the result to zero (the condition for a maximum) gives

$$(21)\qquad \frac{n!}{(c+\tfrac{1}{2})!(n-c-\tfrac{1}{2})!}\,p_1^{\,c+1/2}q_1^{\,n-c-1/2} - \frac{n!}{(c+\tfrac{1}{2})!(n-c-\tfrac{1}{2})!}\,p_2^{\,c+1/2}q_2^{\,n-c-1/2} = 0.$$

Taking logarithms of both expressions in (21) and solving for c yields

$$c = -\frac{1}{2} + \frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1},$$

the same expression as (19), which had been derived exactly. Thus the value of c which maximizes the approximation (20), when rounded, is equal to the c value which maximizes (9). This justifies the use of the approximation and provides an efficient means of dealing with certain optimum problems involving the binomial distribution.

The tables in this paper are based upon the solution developed herein.

SINGLE SAMPLING PLANS FOR PLACING A LOT INTO ONE OF THREE CATEGORIES

Given quality levels p_1, p_2, and $p_3 (p_1 < p_2 < p_3)$ for category 1, category 2, and category 3, respectively, with fixed sample size n, we can write

$$(22)\qquad Pr_{p_1}(1) = \sum_{i=0}^{c_1} \frac{n!}{i!(n-i)!}\,p_1^{\,i}q_1^{\,n-i}$$

$$(23)\qquad Pr_{p_2}(2) = \sum_{i=0}^{c_2} \frac{n!}{i!(n-i)!}\,p_2^{\,i}q_2^{\,n-i} - \sum_{i=0}^{c_1} \frac{n!}{i!(n-i)!}\,p_2^{\,i}q_2^{\,n-i}$$

$$(24)\qquad Pr_{p_3}(3) = 1 - \sum_{i=0}^{c_2} \frac{n!}{i!(n-i)!}\,p_3^{\,i}q_3^{\,n-i}$$

where $Pr_{p_1}(1)$ represents the probability of placing lots of quality p_1 into category 1, $Pr_{p_2}(2)$ represents the probability of placing lots into category 2, and $Pr_{p_3}(3)$ represents the probability of placing lots of quality p_3 into category 3.

Using the criterion discussed in the introductory section, it is desired to maximize the following expression:

$$P = Pr_{p_1}(1) + Pr_{p_2}(2) + Pr_{p_3}(3). \tag{25}$$

It is evident that, for any fixed value of c_2, the solution for c_1 which maximizes (25) is equivalent to the one developed in the preceding section. Similarly, for any fixed value of c_1 the solution for c_2 is equivalent also to the solution developed in the preceding section. Thus the values of c_1 and c_2 which maximize (25) are simply

$$c_1 = -\frac{1}{2} + \frac{n}{\dfrac{\log \dfrac{p_2}{p_1}}{\log \dfrac{q_1}{q_2}} + 1} \quad \text{and} \quad c_2 = -\frac{1}{2} + \frac{n}{\dfrac{\log \dfrac{p_3}{p_2}}{\log \dfrac{q_2}{q_3}} + 1}$$

where c_1 and c_2 are rounded to the nearest integer.

Following this line of reasoning, it is obvious that the preceding results can be applied to the general case of m categories. That is, if $p_1, p_2, \cdots p_m$ are the respective lot qualities of categories 1, 2, 3, $\cdots$, m, then the $c_i (i = 1, 2, 3, \cdots, m-1)$ which maximize

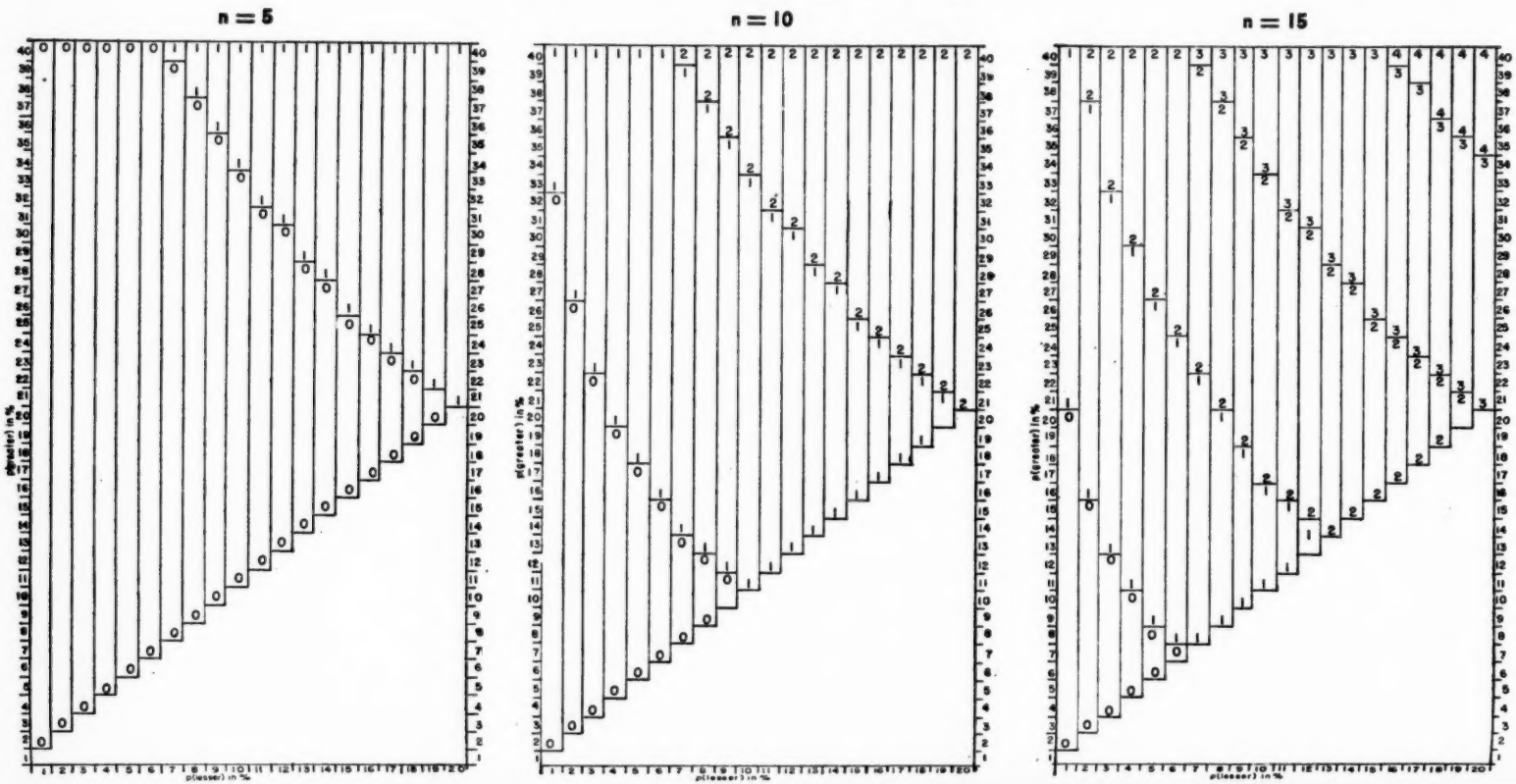
$$P = \sum_{i=1}^{m} Pr_{p_i}(i) \tag{26}$$

are given simply as

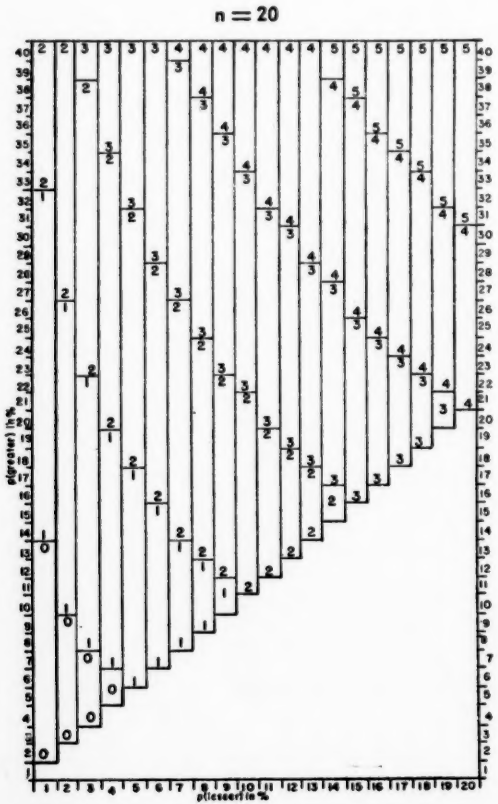
$$c_i = -\frac{1}{2} + \frac{n}{\dfrac{\log \dfrac{p_{i+1}}{p_i}}{\log \dfrac{q_i}{q_{i+1}}} + 1}. \tag{27}$$

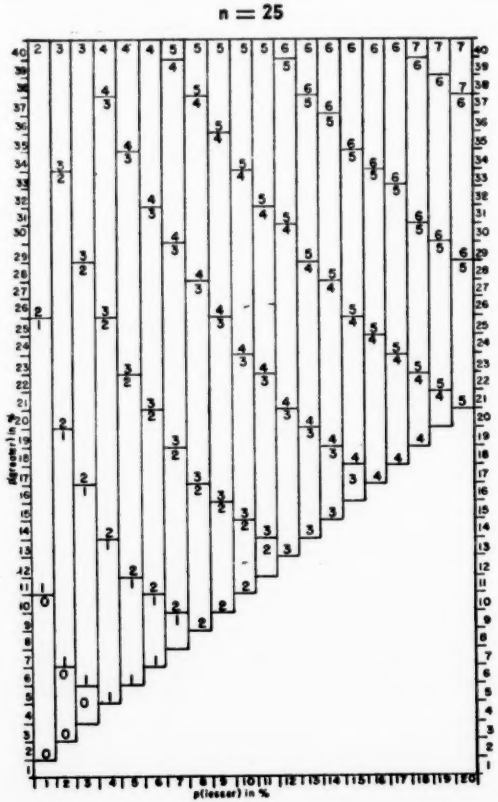
It is obvious from (27) that the tables can be used to obtain any c_i by entering them with p_i and p_{i+1}.

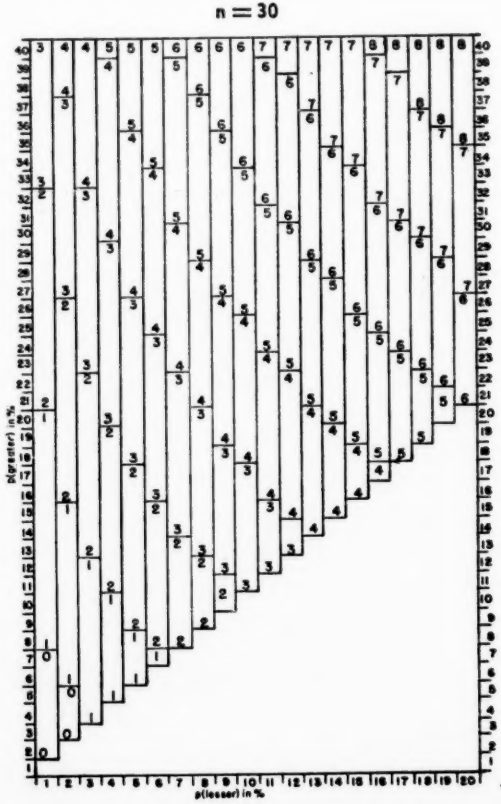
ACCEPTANCE NUMBERS FOR SINGLE SAMPLING PLANS WHICH MINIMIZE THE SUM OF THE PROBABILITIES OF MISCLASSIFICATION



ACCEPTANCE NUMBERS FOR SINGLE SAMPLING PLANS WHICH MINIMIZE THE SUM OF THE PROBABILITIES OF MISCLASSIFICATION







ACCEPTANCE NUMBERS FOR SINGLE SAMPLING PLANS WHICH MINIMIZE THE SUM OF THE PROBABILITIES OF MISCLASSIFICATION

ON ERRORS IN MATRIX INVERSION

PAUL S. DWYER
University of Michigan
AND
FREDERICK V. WAUGH
Bureau of Agricultural Economics

I. INTRODUCTION

MATHEMATICAL statisticians have given a great deal of attention to sampling errors, much less to errors of computation, and very little to errors due to faulty data. Before the statistician has confidence in his results, he should explore all three kinds of errors.

This paper does not consider errors of sampling, since much has been written on that subject. A section is devoted to the discovery of the amount of error resulting from the use of approximate computational methods, but the paper is devoted primarily to a consideration of the effects of errors in the original data—and specifically to the size of the maximum error in a particular element of an inverse matrix due to the inherent errors in the elements of the original matrix.

The computation of an inverse matrix is the key to such important statistical problems as multiple regression, factor analysis, and the Leontief technique of "inter-industry analysis." The methods used in this paper might be applicable to related problems, such as finding bounds to the errors in solving simultaneous equations, but it seemed best to limit this paper to errors in elements of the inverse matrix.

Perhaps a few words should be said here about the nature of inherent errors (or inaccuracies) in the original data. Seldom can a statistician get absolutely accurate data in any field of work. Even in physical experiments, the data are at least subject to errors of measurement. In economics, the statistician must work with estimates of such variables as crop yields, price indices, and national income. In all such cases, it is extremely important that the statistician have information concerning the reliability of the data he uses. Unfortunately few statistical agencies of the Government, or of the trade, provide such information. This situation should be rectified to whatever extent is possible. In any case, a statistician should always find out as much as he can about the data he uses, and should form at least a rough judgment of the maximum errors associated with each series. Without such judgment, he cannot be sure that his analysis is valid.

A recent book [5] and paper [4] by one of the authors give considerable attention to problems related to the subject of the present paper.

A formula for the first order discrepancies in the elements of the inverse matrix due to inherent errors of the data is presented [5, p. 286]. The present paper extends these results

(a) by presenting methods for establishing a bound for the total discrepancy,

(b) by showing how the actual extreme discrepancy for each element of the inverse may be determined and calculated.

The results are applicable to the problem in which the elements of the matrix to be inverted are subject to unspecified but limited errors. Thus each element of the matrix is composed of a known part, a_{ij}, and an unknown part, e_{ij}, with the specification that $|e_{ij}| \leqq \eta_{ij}$. The η_{ij} are the bounds for the errors of the elements.

There are several problems which are closely related to the problem of the errors of the inverse matrix. These are

(a) the problem of the bounds and extreme values of the determinant of the matrix A,

(b) the problem of the bounds and extreme values of the roots of equations which feature A^{-1},

(c) the problem of adjusted equations in which the values of the e_{ij} are known.

The amount of material we found which seemed necessary to an adequate discussion of the topic of errors of matrix inversion was so extensive that we decided to treat this specific problem in this paper. The other problems suggested above, though closely related to the problem of the errors of matrix inversion, are not treated here except as they contribute to the subject of this paper, though several references to important papers dealing with these other problems are given.

II. AN EXAMPLE TO BE ANALYZED

An illustration is next presented in order to clarify the nature of our problem. For this purpose, we have chosen a problem of economics and statistics which was recently discussed by one of the authors [26]. The problem was to determine the least expensive combination of feeds for dairy cows which would meet, or surpass, certain stated requirements. This is a practical application of linear programming. The study concluded that a combination of milo, middlings, gluten, and bran was the least expensive dairy ration meeting the requirements for total digestible nutrients, digestible protein, calcium, phosphorous, and total weight. The analysis of this problem involves the computation of an inverse matrix. We wish to determine the maximum possible error in the elements of this inverse matrix.

The essential data are given in Table I.

TABLE I

The original matrix	0.495	0.275	0.098	0.459	
	0.436	0.449	0.264	0.410	
A	0.395	0.557	1.319	0.385	Absolute
	0.423	0.436	0.467	0.467	row-sums
Maximum inherent errors,	0.0010	0.0005	0.0005	0.0010	0.0030
	0.0010	0.0010	0.0005	0.0010	0.0035
	0.0010	0.0010	0.0015	0.0010	0.0045
	0.0010	0.0010	0.0010	0.0010	0.0040
Computed inverse,	6.483	5.153	3.311	−13.626	28.573
	−4.798	6.485	−0.841	−0.285	12.409
C	0.694	−1.627	1.084	−0.147	3.552
	−2.086	−9.095	−3.299	14.896	29.376
Absolute column-sums	14.061	22.360	8.535	28.954	73.910

The first section of Table I exhibits the original matrix. The elements, a_{ij}, represent the proportion of the jth requirement, which can be supplied by one dollar's worth of the ith feed. These elements were computed using average wholesale prices in Kansas City from October 1949 through September 1950, and using commonly accepted requirements.

The second section exhibits an estimate of the maximum absolute error, η_{ij}, associated with each element of A. These estimates are not exact, and warrant further study, but will serve the purposes of this paper. Obviously, an error of at least 0.0005 is possible for each element, since the data were rounded to three decimals. Somewhat larger errors are possible because of variation in the nutritive content of the feeds. We assume no error in prices. The problem might be stated thus: If these feeds could be bought at quoted prices in the 1949–50 feeding year, and if we assume feeds of approximately the average quality, how much of each of the four feeds would be required? The assumed η_{ij} appear reasonable from this standpoint. We also consider the case in which $\eta_{ij}=\eta=0.0005$. This would be an appropriate assumption if it were assumed that feeds of exactly average quality could be bought at the precise quoted price.

Finally, the last section of Table I shows the inverse of A, which we designate C. It was calculated by a method which first obtains the adjoint of A [27] [5, p. 217]. In this way, we can guarantee that the computed elements of C are correct to the three decimal places recorded. This value of the inverse, the elements of which are correct to three decimal places, can also be computed by the method of Section 13.

To the right of Table I are shown the sums of absolute values of elements of rows. Below are shown the sums of absolute values of elements of columns. These sums are used in our analysis.

We can guarantee that C is computed to three decimals of accuracy. But we have not considered the possible effects of the inherent errors, η_{ij}. The main purpose of this paper is to determine bounds, or extreme values, to the discrepancies between elements of C, and elements of the inverse of a "true" matrix which is free from inherent errors.

III. GENERAL FORMULAS FOR THE INHERENT ERRORS OF THE INVERSE

We first consider the accumulation of the inherent errors. Let T be a real matrix which is composed of an approximate part A and an error part E. Then

$$T = A + E. \tag{3.1}$$

The matrix E is usually composed of unknown elements, e_{ij}, which may be positive, negative, or zero, but which are not greater in absolute value than some bound η_{ij}. We desire to find the inverse of T knowing the value of A and the bounds for E, or alternatively, we desire to know how much the inverse of T might differ from the inverse of A. It is necessary of course that both A and T be non-singular.

In most applied problems the bounds for E are small when compared with the corresponding elements of A, i.e., the relative errors of the elements of A are small. A major concern of this paper is with this situation though some of the formulas are appropriate to the more general case.

In the first part of the paper we assume that A^{-1} can be obtained by exact methods, or, at least, that a correct k decimal place approximation can be computed. We get immediately from (3.1) writing $C = A^{-1}$

$$T^{-1} = (A + E)^{-1} = [(I + EC)A]^{-1} = C(I + EC)^{-1}. \tag{3.2}$$

If the elements of E are small enough so that any norm[1] of EC is less than 1, it can be shown [25] that T^{-1} can be expressed in terms of the convergent series [7, p. 40],

$$T^{-1} = C[I + EC + (EC)^2 - (EC)^3 + \cdots]. \tag{3.3}$$

Pre-multiplication by A shows that

$$I - EC + (EC)^2 - (EC)^3 + \cdots = AT^{-1} = (I + EC)^{-1}. \tag{3.3a}$$

[1] See Section IV for a discussion of norms.

The accumulation of the inherent errors in the inverse is given by the difference between T^{-1} and $A^{-1}=C$. In this paper we call this the discrepancy matrix, D. Then

$$D = T^{-1} - C = -CEC[I - EC + (EC)^2 - \cdots] \tag{3.4}$$

$$D = T^{-1} - C = -C[EC - (EC)^2 + (EC)^3 - \cdots]. \tag{3.4a}$$

Equations (3.4) and (3.4a) are basic to all the sections of this paper dealing with inherent errors.

From (3.1) with $C=A^{-1}$ one has

$$\begin{aligned} D = T^{-1} - A^{-1} &= T^{-1}(A - T)A^{-1} = -T^{-1}EC \\ &= -(D + C)EC \end{aligned} \tag{3.5}$$

so, solving for D

$$D = -CEC(I + EC)^{-1}. \tag{3.6}$$

Alternative forms of the above formulas are

$$T^{-1} = (A + E)^{-1} = [A(I + CE)]^{-1} = (I + CE)^{-1}C \tag{3.2)'}$$

$$T^{-1} = [I - CE + (CE)^2 - (CE)^3 \cdots]C \tag{3.3)'}$$

$$D = T^{-1} - C = -[I - CE + (CE)^2 - \cdots]CEC \tag{3.4)'}$$

$$D = T^{-1} - A^{-1} = -CET^{-1} = -CE(C + D) \tag{3.5)'}$$

$$D = T^{-1} - C = -(I + CE)^{-1}CEC. \tag{3.6)'}$$

Neglecting terms of (3.4) of order higher than one in E we get:

$$D_1 = -CEC, \tag{3.7}$$

where D_1 is the matrix of first-order discrepancies.

This formula could be written somewhat symbolically as

$$d(A^{-1}) = -A^{-1}(dA)A^{-1} \tag{3.8}$$

which is a matrix generalization of the formula for $d(x^{-1})$. Other derivations of this formula (or of the corresponding derivative formula) are available [5, p. 285], [14], [6].

IV. THE NORM OF THE DISCREPANCY MATRIX

We wish to establish an upper bound to the elements of D. One method is to use some norm of D. For our purposes we define $N(A)$, the norm of the real matrix A, as a quantity which satisfies the conditions

(4.1)
(a) $N(A) \geqq 0$
(b) $N(A) = 0$ if and only if $A = 0$
(c) $N(cA) = |c| N(A)$ where c is a scalar
(d) $N(A \pm B) \leqq N(A) + N(B)$
(e) $N(AB) \leqq N(A)N(B)$
(f) $N(A) = 1$ if A is a fundamental unit matrix (with all elements zero except one element which is unity).

These conditions are in substantial agreement with those advocated by Bowker [2]. Application of these conditions to (3.4a) gives with $N(EC) < 1$

$$N(D) \leqq \frac{N(C)N(EC)}{1 - N(EC)} \tag{4.2}$$

and if $N(E)N(C) < 1$,

$$N(D) \leqq \frac{N(E)[N(C)]^2}{1 - N(E)N(C)}. \tag{4.3}$$

Several norms are available. In this section we feature norms advocated recently by Hotelling [11], Bowker [2], and the maximum coefficient, related to a norm, advocated by Turing [24]. These are

$$N_a(A) = \left(\sum_{i,j} a_{ij}^2 \right)^{1/2} = (\text{trace } A'A)^{1/2}, \tag{4.4}$$

$$N_b(A) = \max_i \sum_j |a_{ij}|, \tag{4.5}$$

$$M(A) = \max_{i,j} |a_{ij}|. \tag{4.6}$$

It is to be noted that $nM(A)$ is a norm. The $M(A)$ used by Lonseth [13] appears to be a norm as does the $\max_i [\lambda_i(A'A)]^{1/2}$, where λ_i is the characteristic root of $A'A$, of Wittmeyer [29]. However the $M(A)$ and $m(A)$ of Lonseth and the $\max_i [\lambda_i(A'A)]^{1/2}$ and $\min_i [\lambda_i(A'A)]^{1/2}$ of Wittmeyer play the more precise role of bounds.

We calculate norms for the discrepancy matrix of Table I. We compute $N(E)$ from the η_{ij} matrix and $N(C)$ from the elements of the inverse matrix. We have

(4.7) $N_a(E) \leqq 0.003873$, $N_a(C) = 25.60$ so $N_a(D) \leqq 2.82$.

(4.8) $N_b(E) \leqq 0.0045$, $N_b(C) = 29.376$ so $N_b(D) \leqq 4.47$.

If the Bowker norm is applied columnwise rather than row-wise, we get

(4.9) $N_b'(E) \leqq 0.0044$, $N_b'(C) = 28.954$, so $N_b'(D) \leqq 3.79$.

It can also be shown that when $n^2M(E)M(c)<1$

$$M(D) \leqq \frac{n^2M(E)[M(C)]^2}{1 - n^2M(E)M(C)}. \tag{4.10}$$

Then

(4.11) $M(E) \leqq 0.0015$, $M(C) = 14.896$, so $M(D) \leqq 8.29$.

The use of $nM(A)$ as a norm gives

(4.12) $N_M(E) \leqq 0.0060$, $N_M(C) = 59.584$, so $N_M(D) \leqq 33.154$

and no element of D is larger than $33.154/4 = 8.29$ as is shown in (4.11).

From the calculations above we could use any of 2.82, 4.47, 3.79, or 8.29 as an upper bound to the discrepancy elements. We are free to choose the lowest available bound. Naturally we prefer a bound of 2.82 to one of 8.29 but even the bound 2.82 is too high for much confidence in the precision of the elements of the inverse matrix of Table I. It is not surprising though that the norm which requires the most detailed calculation is the one which gives us the best results.

These norms are all relatively easy to compute and in some cases, particularly when all the terms of the matrix have approximately equal absolute values, may be satisfactory. In general, however, they provide rather loose bounds to the discrepancies, d_{ij}. Commonly, as in the present case, we need a closer bound for each d_{ij}. We proceed therefore to consider the possibility of establishing closer bounds, or perhaps lower extreme values, than those found in this section of the paper.

V. BOUNDS FOR THE ELEMENTS OF THE DISCREPANCY MATRIX WHEN THE ERRORS HAVE A COMMON BOUND

An important special case is that in which the maximum absolute error in any element of A is η. For example, if the elements of A are accurate to three decimal places, the value of the error of any element is between -0.0005 and 0.0005, and we say $\eta = 0.0005$.

A method is available [5, p. 285] for computing the first order approximations to the elements of the discrepancy matrix. This is done by assigning the elements of E in (3.9) as either $-\eta$ or η in such a manner as to make a particular element of D_1 take on the extreme value.

A different set of values of E may be necessary to obtain the extreme value for each term in the first order discrepancy matrix. The results are given by the formula

$$B[d_{1.rs}] = \eta \sum_j |c_{rj}| \sum_i |c_{is}| . \tag{5.1}$$

The calculations are illustrated in Table II where the A and C matrices are those of Table I and $\eta = 0.0005$. The sums of the absolute values of the rows and columns of c_{ij} are placed in the margins of the first matrix of Table II. These marginal values are multiplied to obtain the elements of the matrix. These products are multiplied by $\eta = 0.0005$ to obtain the values $B[d_{1;rs}]$ which are recorded in the second matrix.

TABLE II

	Products				Row Sum
Products	401.765	638.892	243.871	827.303	28.573
	174.483	277.465	105.911	359.290	12.409
	49.945	79.423	30.316	102.845	3.552
	413.056	656.847	250.724	850.553	29.376
Column sum	14.061	22.360	8.535	28.954	73.910
$B[d_{1;rs}]$	0.201	0.319	0.122	0.414	
	0.087	0.139	0.053	0.180	
	0.025	0.040	0.015	0.051	
	0.207	0.328	0.125	0.425	
$B[d_{rs}]$	0.209	0.331	0.127	0.430	
	0.090	0.144	0.055	0.187	
	0.026	0.042	0.016	0.053	
	0.215	0.341	0.130	0.441	

How good are these first order discrepancies? One of the purposes of this paper is to show that the results of the method described above can be extended, with little more computational work, to provide bounds for the actual discrepancy matrix rather than for bounds for the first order approximation to it. This can be done if $N(EC) < 1$ and if $\eta < 1/(\sum_i \sum_j |c_{ij}|)$.

We show below that a bound for d_{rs} is given by the formula

$$B[d_{rs}] = \frac{\eta \sum_j |c_{rj}| \sum_i |c_{is}|}{1 - \eta \sum_i \sum_j |c_{ij}|} . \tag{5.2}$$

This formula was given essentially, though not explicitly, by Lonseth [13, p. 197] who directed his attention primarily to bounds for the errors of solutions of equations rather than to bounds for the errors of the inverse matrix. The general nature of Lonseth's proofs should be of interest to the more mathematically inclined readers.

This formula indicates that we need only multiply the first order discrepancies by $1/(1-\eta \sum_i \sum_j |c_{ij}|)$ to get discrepancy bounds. Alternatively these bounds may be obtained by multiplying the products in Table II by $\eta/(1-\eta \sum_i \sum_j |c_{ij}|)$. Thus the values in the last matrix of Table II are obtained by dividing the values of $B[d_{1;rs}]$ by 0.963045. Slightly more accurate values, since they avoid the rounding off errors of the $B[d_{1;rs}]$ terms, may be obtained by multiplying the elements of the first matrix of Table II by 0.000519187.

In this case the first order bounds underestimate the bounds by about 3.84 per cent of the first order bounds. More generally the adequacy of the first order bounds may be measured by the formula

$$\frac{B[d_{rs}]}{B[d_{1;rs}]} - 1 = \frac{1}{1 - \eta \sum_i \sum_j |c_{ij}|} - 1. \tag{5.3}$$

The first order bounds are very unsatisfactory if $\eta \sum_i \sum_j |c_{ij}|$ is nearly one and they are of doubtful value unless $\eta \sum_i \sum_j |c_{ij}| < 0.1$. Equation (5.3) implies that η should be less than $1/(10 \sum_i \sum_j |c_{ij}|)$ if the first order bounds are to underestimate the bounds by 11% or less of the first order bounds.

We now prove (5.2). If η is the maximum absolute value of any element of E, then the maximum absolute value of each element of EC is given by

$$B[EC] = \begin{bmatrix} \eta \sum |c_{i1}| & \eta \sum |c_{i2}| & \cdots & \eta \sum |c_{in}| \\ \eta \sum |c_{i1}| & \eta \sum |c_{i2}| & \cdots & \eta \sum |c_{in}| \\ \cdots & \cdots & \cdots & \cdots \\ \eta \sum |c_{i1}| & \eta \sum |c_{i2}| & \cdots & \eta \sum |c_{in}| \end{bmatrix}. \tag{5.4}$$

We note that the rows of $B[EC]$ are identical. This is true also for the rows of $B[EC]^2$, $B[EC]^3$, etc. Now $B[EC]^2$ may be found by multiplying each element of $B[EC]$ by $\eta \sum_i \sum_j |c_{ij}|$; $B[EC]^3$ may be found by multiplying each element of $B[EC]$ by $[\eta \sum_i \sum_j |c_{ij}|]^2$, etc. Finally if $\eta \sum_i \sum_j |c_{ij}| < 1$ and if we put

$$k = 1 - \eta \sum_i \sum_j |c_{ij}| \tag{5.5}$$

we have

$$(5.6)\quad B\left[\sum_{p=1}^{\infty}(-1)^p(EC)^p\right] \leqq \begin{bmatrix} \frac{\eta\sum|c_{i1}|}{k} & \frac{\eta\sum|c_{i2}|}{k} & \cdots & \frac{\eta\sum|c_{in}|}{k} \\ \frac{\eta\sum|c_{i1}|}{k} & \frac{\eta\sum|c_{i2}|}{k} & \cdots & \frac{\eta\sum|c_{in}|}{k} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\eta\sum|c_{i1}|}{k} & \frac{\eta\sum|c_{i2}|}{k} & \cdots & \frac{\eta\sum|c_{in}|}{k} \end{bmatrix}$$

and by taking bounds in (3.4a) we get (5.2).

Table II provides bounds to d_{ij} on the assumption that the maximum error in each element of A is $\eta = 0.0005$. Section IV of this paper provides norms for the case in which η_{ij} varies from 0.0005 to 0.0015. We could use the methods of this section, with $\eta = 0.0015$ and find the bound for the term having the largest possible error. This is at the junction of the row and column of C having the largest absolute sums. Thus from Table I, we have at once

$$B[d_{44}] = \frac{(0.0015)(28.954)(29.376)}{1 - (0.0015)(73.910)} = 1.435$$

which, though obviously larger than the value of 0.441 in Table II with $\eta = 0.0005$, is appreciably smaller than the bounds established in Section IV with the use of norms. The methods of Section V are almost as easy computationally as those of Section IV and often, as in this case, give much more satisfactory results.

The bounds above are quite general, applying to any inverse matrix and to any assumed error bounds, η_{ij}. For example, the results to this point show that the absolute error in c_{44} of Table I cannot be greater than 0.441 if we assume that each $\eta_{ij} = 0.0005$, nor can it be greater than 1.435 if the maximum η_{ij} are as stated in the second section of Table I. Bounds of this kind can be computed quickly and easily, and often should be close enough for our purposes.

But, in general, it is possible to reduce the bounds in two ways: First, by taking account of the signs of error terms; and, second, by using the complete error matrix, η_{ij}, when its elements are unequal. Later sections of this paper determine the lowest possible bounds for the general case; —that is, they show the extreme attainable values of d_{rs}, taking account of signs of error terms, and making use of the entire η_{ij} matrix.

Before doing this, however, we shall briefly consider a special class of matrices in which there is no problem of signs.

VI. A SPECIAL CASE IN WHICH $B[d_{rs}]$ IS AN ATTAINABLE EXTREME

There are certain special cases in which the bounds of Section V are actually attainable, and thus are the best possible bounds when $\eta_{ij}=\eta$. One of these is the case in which all elements of C have the same sign, i.e., the signs are all positive or all negative.

The case in which the signs of C are all positive is of current interest, since it includes the so-called "Leontief Matrices" used in "inter-industry analysis" of "input-output studies." Recent proofs are available [25, p. 150] and [9] that the inverse of the Leontief matrix is composed entirely of non-negative elements.

The positive nature of the elements of the inverses of matrices having the mathematical properties of "Leontief" matrices was established in 1907 by Frobenius [8] and in 1920 by Janet [12, pp. 135–41]. Ostrowski [18] [19] developed the properties of these matrices in considerable detail. Though Ostrowski directed his efforts toward the determination of the bounds for the determinants, he also gave excellent surveys of previous work including the allied topic of bounds for the errors of the solutions of linear equations.

More recently Taussky [23] has written on these determinants and Woodbury [31] has written on the properties of "Leontief" matrices. They provide extensive bibliographies. The stability of the inverse Leontief matrix has been studied by Woodbury [31].

We first illustrate the fact that if the inherent errors associated with the elements of a matrix have a common bound, η, the methods of Section V provide the lowest possible bound to the discrepancy terms of the inverse. In other words, in the Leontief case $B[d_{rs}]$ is an attainable extreme.

In actual practice, inter-industry analyses are usually made by the use of Leontief matrices of large order. The Bureau of Labor Statistics has constructed such a matrix of order 450, and plans to invert one of order 200. However, the principle can be shown by considering the simple 3-rowed Leontief matrix in Table III.

We note that the computed inverse, C, is composed entirely of positive elements. The discrepancy bounds $B[d_{rs}]$ are computed by (5.2) using the row-sums and column-sums of C, and assuming $\eta=0.005$. These discrepancies could be actually reached in the extreme case in which the error of each element of A is negative. This is shown in the final two sections of the table. First we adjust A by subtracting 0.005

from each element, then invert the adjusted matrix. The differences between the inverse of the adjusted matrix and the inverse of the original matrix are precisely $B[d_{rs}]$ as computed, aside from an error of one digit in the first element of the second row and the second element of the third row. These errors are due to rounding of both inverses to two decimals.

The principle illustrated with this simple numerical matrix holds for any Leontief matrix, no matter how large, as is shown in Section VIII. The $B[d_{rs}]$ bound provides extreme values of the discrepancies which would be reached if all inherent errors were at their maxima, and all were negative. Therefore, the bound cannot be improved in case of a Leontief matrix with a common bound to the inherent errors.

TABLE III

Leontief matrix, $A =$	1.00	−0.50	−0.30	
	−0.40	1.00	−0.30	
	−0.20	−0.30	1.00	
				Absolute row-sums
Computed inverse, $C =$	1.56	1.01	0.77	3.34
	0.79	1.61	0.72	3.12
	0.55	0.68	1.37	2.60
Absolute column-sums	2.90	3.30	2.86	$\sum\sum\lvert a_{ij}\rvert = 9.06$
Discrepancy bounds, assuming $\eta = 0.005$, $B[d_{rs}] =$	0.05	0.06	0.05	
	0.05	0.05	0.05	
	0.04	0.04	0.04	
Adjusted matrix $B = [a_{ij} - 0.005] =$	0.995	−0.505	−0.305	
	−0.405	0.995	−0.305	
	−0.205	−0.305	0.995	
Computed inverse of B	1.61	1.07	0.82	
	0.83	1.66	0.77	
	0.59	0.73	1.41	

VII. A LEONTIEF MATRIX WITH UNIFORM PERCENTAGE ERRORS

In some practical cases, the bound of an inherent error η_{ij} may be some proportion of a_{ij}. For example, we may set the bound

$$\eta_{ij} \leqq k a_{ij} \qquad 0 < k < 1. \tag{7.1}$$

Consider a Leontief matrix, with inherent errors bounded by (7.1). In this case we can determine a very simple bound to the discrepancies in the inverse. Since each element of the inverse is positive, the extreme discrepancies will occur when all errors are negative. In this case, the "true" matrix would be

$$(7.2) \qquad \begin{bmatrix} a_{11}(1-k) & a_{12}(1-k) & \cdots & a_{1n}(1-k) \\ a_{21}(1-k) & a_{22}(1-k) & \cdots & a_{2n}(1-k) \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}(1-k) & a_{n2}(1-k) & \cdots & a_{nn}(1-k) \end{bmatrix}$$

and more briefly,

$$(7.3) \qquad T = A(I - kI).$$

The inverse of the true matrix would be

$$(7.4) \qquad T^{-1} = \frac{I}{1-k} A^{-1} = \frac{I}{1-k} C,$$

which would be obtained by multiplying each element of C by $1/(1-k)$. Therefore, the maximum discrepancy in this case is

$$(7.5) \qquad B[d_{rs}] = \frac{C_{rs}}{1-k} - C_{rs} = \frac{k}{1-k} C_{rs}.$$

Thus, if the maximum inherent error associated with any element of a Leontief matrix is 5 per cent of that element, the maximum discrepancy associated with any element of the inverse is 0.05/0.95 times that element of the inverse. This is the closest bound which could be established, since the discrepancies are extreme values which could be reached if the errors were distributed in the worst possible way. The bound stated in (7.5) holds for a Leontief matrix of any order and is extremely simple to apply.

VIII. LEMMAS CONCERNING THE EXTREME DISCREPANCIES

The discussion of the earlier sections is concerned with bounds for the first order discrepancy matrix and the discrepancy matrix. Aside from the methods of Section VI and Section VII, which are applicable to special matrices, there is no guarantee that the bound can actually be attained. In this and the sections of the paper immediately following we work out somewhat more complicated techniques for finding the best possible bounds, i.e., smaller bounds than those described above and bounds which cannot be replaced by yet smaller bounds because these can actually be attained in some extreme problem. These discrepancy bounds we call "extreme discrepancies." The theory also enables us to produce the matrix, $A+E$, which would maximize the discrepancy between a particular element of A^{-1} and the corresponding element of

$(A+E)^{-1}$. The derivation is somewhat more involved than is the application, which is very satisfactory for the case of common bounds.

The solution is based on two lemmas, proved below, which establish the fact that a specific sign can be assigned to each element of the error matrix, once the c_{ij} are known, so as to obtain the extreme discrepancy for any given element of the inverse matrix. The first lemma deals with extreme first order discrepancies and the second with extreme total discrepancies. With the signs determined, it is possible to find the discrepancy matrix which results from the specified errors in several ways.

We know from (3.9) that the first order approximation to the discrepancy matrix is given by $-CEC$. Formal expansion of this product shows that the value of the element in the rth row and sth column of the first order discrepancy matrix is $-C_rEC_s$ where C_r is the vector of elements in the rth row and C_s is the vector of elements in the sth column. The first order discrepancy for the element in the rth row and sth column of the inverse can then be written

$$d_{1;rs} = -\sum_i \sum_j c_{ri}e_{ij}c_{js}. \tag{8.1}$$

We are free to give the signs to the e_{ij} so as to obtain the extreme $d_{1;rs}$. If we take the sign of e_{ij} as opposite to that of the product $c_{ri}c_{js}$, then $d_{1;rs}$ is as large (positively) as is possible. Similarly if the sign of e_{ij} is taken as that of $c_{ri}c_{js}$, $d_{1;rs}$ takes on its extreme negative value. We introduce the definitions

$$\begin{aligned}
&\theta_{ij} = 1 \quad \text{when } c_{ri} \text{ has the same sign as } c_{js} \\
&\theta_{ij} = -1 \quad \text{when } c_{ri} \text{ has the opposite sign to that of } c_{js} \\
&\theta_{ij} = 0 \quad \text{when } c_{ri} \text{ or } c_{sj} = 0, \\
&e_{ij} = \theta_{ij}\epsilon_{ij} \quad \text{where } \epsilon_{ij} \geqq 0.
\end{aligned} \tag{8.2}$$

Then (8.1) becomes

$$\underline{L}[d_{1;rs}] = -\sum_i \sum_j c_{ri}\theta_{ij}\epsilon_{ij}c_{js} \tag{8.3}$$

where $\underline{L}[d_{1;rs}]$ indicates the smallest possible (attainable) value of $d_{1;rs}$. Similarly if $\theta_i'_j = -\theta_{ij}$, we have

$$\overline{L}[d_{1;rs}] = \sum_i \sum_j c_{ri}\theta_{ij}'\epsilon_{ij}c_{js} = \sum_i \sum_j c_{ri}\theta_{ij}c_{js} \tag{8.4}$$

is the largest positive value of $d_{1;rs}$.

The extreme first order discrepancy of a given element is not necessarily associated with the extreme first order discrepancy of another

element. Theoretically it is necessary to fix the signs of each e_{ij} in order to obtain the extreme discrepancy for some one element. Sometimes these signs will also maximize the discrepancies associated with several other elements, or even of all elements. Thus if the values of c_{ij} are all positive, all $\theta_{ij}=1$, $\theta_{ij}'=-1$ no matter what the values of r and s. If the e_{ij} take on values as large as η_{ij} in absolute value, (8.3) and (8.4) become

$$\begin{aligned} \underline{L}[d_{1:rs}] &= -\sum_i \sum_j c_{ri}\theta_{ij}\eta_{ij}c_{js} \\ \overline{L}[d_{1:rs}[&= \sum_i \sum_j c_{ri}\theta_{ij}\eta_{ij}c_{js}, \end{aligned} \tag{8.5}$$

while if $\eta_{ij}=\eta$, we get

$$\begin{aligned} \underline{L}[d_{1:rs}] &= -\eta\sum_i \sum_j c_{ri}\theta_{ij}c_{js} \\ \overline{L}[d_{1:rs}] &= \eta \sum_i \sum_j c_{ri}\theta_{ij}c_{js}. \end{aligned} \tag{8.6}$$

We now state Lemma 1. *The extreme first order discrepancies of c_{rs} are given by (8.5) and (8.6).*

Once the signs of the e_{ij} are determined, the values of the other elements of the first order discrepancy matrix, though not necessarily extreme discrepancies, may be computed with the formulas

$$\begin{aligned} d_{1:pq} &= \mp \sum_i \sum_j c_{pi}\theta_{ij}\eta_{ij}c_{jq} \\ d_{1:pq} &= \mp \eta\sum_i \sum_j c_{pi}\theta_{ij}c_{jq}. \end{aligned} \tag{8.7}$$

We next consider the question of finding extreme discrepancies rather than extreme first order discrepancies. Using (3.5) and (3.5)′ we get

$$D = T^{-1} - C = T^{-1}EC = -CET^{-1}. \tag{8.8}$$

If we knew the signs of the elements of T^{-1} we could determine the signs of E so as to obtain an extreme value of d_{rs}, just as we did of $d_{1;rs}$ above. We do not know the value of T^{-1}, but if the relative errors are small, the elements of T^{-1} are close to those of C. *If the relative errors are small enough so that the signs of the elements of T^{-1} are identical with the signs of the elements of C, then the extreme discrepancies occur simultaneously with the extreme first order discrepancies.* This is Lemma 2.

One does not know, of course, at the beginning that this condition of identical signs is satisfied. But he computes the greatest possible discrepancy for each element, with the given bounds for e_{ij}, and knows

that he has the extreme discrepancy for each element if the discrepancy is less in absolute value than the corresponding element of C.

The calculation of this potential extreme discrepancy is accomplished in various ways, as is shown in the sections following. We wish to indicate here that the value of this extreme discrepancy element

$$L[d_{rs}] = -L[d_{1:rs}] \cdot [I - EC + (EC)^2 - \cdots] \tag{8.9}$$

is obtained by using the same E in the series which is used in obtaining $L[d_{1:rs}]$ from CEC. This formula can also be written as

$$D = -CEC + CECEC - CECECEC + \cdots \tag{8.10}$$

where E is taken so as to obtain extreme discrepancies for d_{rs}.

We now return to the problem discussed in Section VI. We know that the elements of the inverse matrix are all non-negative so the signs of all e_{ij} terms are taken as non-negative. The extreme negative first order discrepancies are obtained by setting $e_{ij} = \eta_{ij}$. When $\eta_{ij} = \eta$, the absolute values of the first order discrepancies become $\eta \sum_{i,j} c_{ri}\theta_{ij}c_{js}$. Since the elements of θ and those of the inverse matrix are non-negative, this becomes $\eta \sum_i c_{ri} \sum_j c_{js}$ and this is the equivalent of (5.1). Hence (5.1) gives the absolute value of the extreme first order discrepancies.

If all the extreme first order discrepancies are smaller than the corresponding values of c, the formula (5.2) gives the extreme discrepancies, as is illustrated in Table III and as clarified in sections following which treat also the case of unequal bounds.

The Leontief matrix is not the only special case in which all maximum discrepancies are attained at once. Any matrix having an inverse whose elements alternate in sign also has this property. More generally, if the signs of each row (column) of the inverse matrix are equal to (or the negative of) the signs of every other row (column), the matrix takes on extreme values, positive or negative, simultaneously. Thus the elements of the matrix having an inverse with signs

$$\begin{array}{cccc} - & + & + & - \\ - & + & + & - \\ + & - & - & + \\ + & - & - & + \end{array}$$

take on extreme values simultaneously.

Some elements may take on extreme values simultaneously even though not all the elements do, if the signs of the corresponding rows (or columns) of the inverse matrix are identical or opposite. Thus the

extreme negative discrepancy of a_{14} in Table I occurs simultaneously with the extreme positive discrepancy of a_{44}.

IX. THE COMPUTATION OF THE INVERSE OF AN ADJUSTED MATRIX

Once we know just what are the values of E which give extreme discrepancies, we can avoid the expansion of a series suggested in the last section by using techniques which have been worked out to handle adjusted matrices [5, p. 292]. An adjusted matrix is one in which some element (or elements) is replaced by some other element (or elements). Formulas featuring determinants were worked out by Sherman and Morrison [21] [22] and Woodbury [15] for the case of adjustment of a single element and for the case of the adjustment of the elements of a column or row. These adjustments have been described in matrix form [5, p. 293]. A more general type of adjustment has been worked out by Bartlett [1] in which all elements of the matrix may be adjusted if the matrix of adjustments can be expressed as the product of two vectors. Still more general formulas for the adjustment of matrices have been given by Woodbury [30]. The attention of the reader is also called to the earlier (1936) work of Wittmeyer [29] who used $\max_i [\lambda_i(A'A)]^{1/2}$ and $\min_i [\lambda_i(A'A)]^{1/2}$ in obtaining bounds for the errors in adjusted systems of equations, and to the recent work of engineers [28] [16] [3].

The two general formulas given by Woodbury, which are slightly altered here to parallel the formulas of Section 3, are

$$(9.1) \qquad D = (A + USV)^{-1} - A^{-1} = -CUS(S + SVCUS)^{-1}SVC$$

or if S is non-singular,

$$(9.2) \qquad D = (A - US^{-1}V)^{-1} - A^{-1} = CU(S - VCU)^{-1}VC.$$

These are too general to be useful to us here in their general form, though special cases are very useful. For the general problem it might be more practical to calculate $(A+E)^{-1}$, when E is known, by the method of the next section rather than to follow the computation of $(S-VCU)^{-1}$, or of $(S+SVCUS)^{-1}$, by the additional matrix operations demanded by (9.1) or (9.2).

These formulas are very useful for the special case in which S is a scalar, say σ. Then u is a row vector, v a column vector and

$$(9.3) \qquad D = (A + u\sigma v)^{-1} - A^{-1} = -Cu\sigma(\sigma + \sigma vCu\sigma)^{-1}\sigma vC$$

becomes

$$(9.4) \qquad D = (A + u\sigma v)^{-1} - A^{-1} = -\frac{\sigma CuvC}{1 + \sigma vCu}.$$

Proof of (9.4) follows by pre-multiplication or post-multiplication by $A+E=A+u\sigma v$. Following the style of Bartlett [1], one may arrive at this formula by means of formal substitution in (3.4).

This formula for the inverse of an adjusted matrix is very useful in obtaining the extreme discrepancy matrix when the errors of the elements of the original matrix have common bounds, as is shown in Section XI.

X. THE RECALCULATION OF THE INVERSE

The equation (3.2) rather than (3.4) might be used once the values of the error matrix which give extreme discrepancies are known. This method calls for the calculation of $T^{-1}=(A+E)^{-1}$ and thence $D=T^{-1}-C$. It is extremely simple in theory and is especially useful where machines are available for inverting matrices efficiently. Ideally the reduction method should be some exact method such as the method of determinants [5, p. 76] though good approximation methods would be satisfactory, particularly if corrected to a given number of decimal places by the method of Section XIII. It may be necessary to make a recalculation for each element of the inverse for which it is desired to obtain the extreme discrepancy.

XI. EXTREME DISCREPANCIES FOR THE CASE OF COMMON BOUNDS

We are now in a position to improve the treatment of the case discussed in Section 5 in that we can produce extreme discrepancies. We have three different means of calculation at our disposal. These are

(a) the method of expansion in series,
(b) the method of adjusted matrices,
(c) the method of recalculation.

We also discuss (d) the treatment of joint extreme elements. We discuss them in order. Application is made to the problem of Tables I and II with $\eta=0.0005$.

(*a*) *The method of expansion in series.* In this case we use the matrix expansion (8.10) and compute the matrices CEC, $CECEC$, $CECECEC$, $\cdots$, until convergence is attained to the desired number of places. If $\eta_{ij}=\eta$ and $E=\eta\theta'$, the value of D is

$$D = \eta[C\theta'C] + \eta^2[C\theta'C\theta'C] + \eta^3[C\theta'C\theta'C\theta'C]\cdots. \tag{11.1}$$

The illustration of this method is presented in Table IV, where the extreme discrepancy for C_{44} and the other associated discrepancies are computed. The value of the matrix θ' is obtained by premultiplying the

column vector of signs of the fourth column of C by the row vector of signs of the fourth row of C. This result is placed at the top of Table IV. The three decimal place value of C is placed below θ' below which appears $\theta'C$. The values $C\theta'C$, $C\theta'C\theta'C$, $C\theta'C\theta'C\theta'C$ are computed and recorded below. The corresponding elements of these matrices are then multiplied respectively by η, η^2, η^3 and the results are added to obtain the discrepancy matrix. The elements of this discrepancy matrix are less in absolute value than the elements of the inverse matrix above so we may say that extreme discrepancy of the d_{44} term is 0.438. This is slightly less than the 0.441 obtained in Table II. If η were larger, the improvement would be greater.

For this problem the d_{44} term is somewhat smaller in absolute value when the values of θ_{ij} are used in place of θ_{ij}'. Thus if $E = -\eta\theta'$ instead of $\eta\theta'$, (11.1) becomes

$$D = -\eta C\theta'C + (-\eta)^2 C\theta'C\theta'C + (-\eta)^2 C\theta'C\theta'C\theta'C + \cdots \tag{11.2}$$

and the values are those of sums of alternating series. Thus in Table IV

$$d_{44} = -0.0005(850.553) + (0.0005)^2(50504) - (0.0005)^3(300.10^4) + \cdots = -0.413.$$

(*b*) *The method of adjusted matrices.*

The value of E must be taken, in order to obtain the extreme r, s term in the discrepancy matrix as

$$E = [e_{ij}] = u\eta v$$

where u is a column vector having elements with absolute value unity with signs of c_r, v is a row vector with unit elements, aside from sign, and with signs identical with those of c_s. The sign of the scalar η is not yet determined. In order to attain the extreme value, positive or negative, the sign of η should be taken opposite to that of vCu so that the denominator of (9.4) becomes as small as possible. If vCu is positive, η should have a negative sign and the formula for the discrepancy matrix is

$$D = -\eta \frac{CuvC}{1 - \eta vCu} = \eta \frac{C\theta'C}{1 - \eta vCu}. \tag{11.3}$$

Now since the numerator of (11.3) is the value of D_1 as found in Section VIII, we need only calculate $1 - \eta vCu$ and divide it into D_1 to find the values of the discrepancy matrix associated with some extreme dis-

TABLE IV

METHOD OF EXPANSION IN SERIES

θ'	1	1	1	−1
	1	1	1	−1
	1	1	1	−1
	−1	−1	−1	1
C	6.483	5.153	3.311	−13.626
	−4.798	6.485	−0.841	−0.285
	0.694	−1.627	1.084	−0.147
	−2.086	−9.095	−3.299	14.896
$\theta'C$	4.465	19.106	6.853	−28.954
	4.465	19.106	6.853	−28.954
	4.465	19.106	6.853	−28.954
	−4.465	−19.106	−6.853	28.954
$C\theta'C$	127.578	545.916	195.811	−827.303
	5.050	21.609	7.751	−32.747
	1.331	5.694	2.042	−8.628
	−131.164	−561.258	−201.314	850.553
$C\theta'C\theta'C$	7,575	32,415	11,627	−49,124
	300	1,283	460	−1,944
	79	338	121	−512
	−7,788	−33,326	−11,954	50,504
$C\theta'C\theta'C\theta'C \cdot 10^{-4}$	45	192	69	−292
	2	8	3	−12
	0	2	1	−3
	−46	−198	−71	300
D	0.066	0.281	0.101	−0.426
	0.003	0.011	0.004	−0.017
	0.001	0.003	0.001	−0.004
	−0.068	−0.289	−0.104	0.438

crepancy element. The method is illustrated in Table V where, again, it is desired to obtain the extreme value of the d_{44} term. The value of C is recorded and the column sums are recorded after each element has been multiplied by the sign of the corresponding element in the fourth column. Corresponding row sums are obtained after the elements have been multiplied by the signs in the fourth row. These marginal entries are summed after multiplication by the signs of the fourth row (or fourth column) to get $vCu = 59.378$. Since $vCu > 0$, η is given a negative

sign. The value of $C\theta'C$ is then obtained by multiplying the marginal entries, as in Table II. The results agree with the results of the more formal methods of Table IV.

The values of D_1 are then obtained by multiplication by 0.0005 and the values of D by dividing the values of D_1 by $1-(0.0005)(59.378) = 0.970311$. The resulting values are in complete agreement with the results of Table IV. They are obtained much more easily with the method of Table V.

We should note further that the other set of extreme discrepancies (having signs opposite to those of this set) are obtained by using the multiplier

$$\frac{-\eta}{1+\eta vCu} = \frac{-0.0005}{1+(0.0005)(59.378)} = \frac{-0.0005}{1.029689}.$$

The value of the extreme negative d_{44} term, for example, is

$$\frac{(-850.553)(0.0005)}{1.029689} = -0.413,$$

as shown above.

TABLE V

METHOD OF ADJUSTED MATRICES

	6.483	5.153	3.311	−13.626	−28.573
	−4.798	6.485	−0.841	−0.285	−1.131
C	0.694	−1.627	1.084	−0.147	−0.298
	−2.086	−9.095	−3.299	14.896	29.376
	−4.465	−19.106	−6.853	28.954	59.378
	127.578	545.916	195.811	−827.303	
$C\theta'C$	5.050	21.609	7.751	−32.747	
	1.331	5.694	2.042	−8.628	
	−131.164	−561.258	−201.314	850.553	
	0.064	0.273	0.098	−0.414	
D_1	0.003	0.011	0.004	−0.016	
	0.001	0.003	0.001	−0.004	
	−0.066	−0.281	−0.101	0.425	
	0.066	0.281	0.101	−0.426	
D	0.003	0.011	0.004	−0.017	
	0.001	0.003	0.001	−0.004	
	−0.068	−0.289	−0.104	0.438	

(c) *The method of recalculation.* The method of recalculation, described in Section X, is illustrated in Table VI. The matrix of error bounds, with signs assigned so as to maximize the d_{44} term, is first exhibited. This matrix is then added to A to obtain the specific $A+E$. The inverse is then computed. This computation is not exhibited in Table VI but the actual calculation used the method of determinants so that the answers can be guaranteed to four decimal places. The discrepancy matrix, formed by subtraction of the four place value of A^{-1}, is recorded to three decimal places. It is identical with the three-place discrepancy matrix of Table IV and Table V.

TABLE VI

METHOD OF RECALCULATION

	−0.0005	−0.0005	−0.0005	0.0005
E	−0.0005	−0.0005	−0.0005	0.0005
	−0.0005	−0.0005	−0.0005	0.0005
	0.0005	0.0005	0.0005	−0.0005
	0.4945	0.2745	0.0975	0.4595
$A+E$	0.4355	0.4485	0.2635	0.4105
	0.3945	0.5565	1.3185	0.3855
	0.4235	0.4365	0.4675	0.4665
	6.5488	5.4340	3.4122	−14.0519
$(A+E)^{-1}$	−4.7956	6.4966	−0.8369	−0.3015
	0.6944	−1.6243	1.0856	−0.1518
	−2.1539	−9.3842	−3.4025	15.3346
	6.4831	5.1527	3.3112	−13 6256
$C=A^{-1}$	−4.7982	6.4853	−0.8408	−0.2846
	0.6937	−1.6270	1.0845	−0.1474
	−2.0863	−9.0950	−3.2987	14.8963
	0.066	0.281	0.101	−0.426
D	0.003	0.011	0.004	−0.017
	0.001	0.003	0.001	−0.004
	−0.068	−0.289	−0.104	0.438

The extreme negative value for d_{44} is obtained by this method by taking the signs of E opposite to those of Table VI. Another recalculation is then necessary.

(d) *Extreme discrepancies associated with the elements.* Frequently we desire the extreme discrepancy associated with each individual element,

rather than the extreme discrepancy of some particular element. This can be done, for the case of common bounds, by the method of adjusted matrices but the values of the extreme $C\theta C$ and $1-\eta vCu$ must be calculated for each element unless we have the fortunate situation, characterized by Leontief matrices, where the extreme discrepancies of many elements are taken on simultaneously, as illustrated in Table III.

The technique is illustrated in the general case in Table VII where $\eta = 0.0005$. The columns of C are added after multiplication by the

TABLE VII

TREATMENT OF JOINT EXTREME ELEMENTS

6.483	5.153	3.311	−13.626	28.573	8.985	18.267	−28.573	28.573
−4.798	6.485	−0.841	−0.285	1.131	12.409	−11.839	−1.131	12.409
0.694	−1.627	1.084	−0.147	0.298	−3.258	3.552	−0.298	3.552
−2.086	−9.095	−3.299	14.896	−29.376	−18.606	−11.186	29.376	29.376
14.061	6.136	8.535	−28.384	57.116	11.924	44.844	−57.116	
3.077	22.360	4.685	−28.660	58.782	43.258	14.062	−58.782	
14.061	6.136	8.535	−28.384	57.116	11.924	44.844	−57.116	
−4.465	−19.106	−6.853	28.954	−59.378	−36.742	−21.166	59.378	
14.061	22.360	8.535	28.954					73.910
401.765	638.892	243.871	827.303	0.971442	0.970609	0.971442	−0.970311	
174.483	277.465	105.911	359.290	0.994038	0.978371	0.994038	−0.981629	
49.945	79.423	30.316	102.845	0.977578	0.992969	0.977578	−0.989417	
413.056	656.847	250.724	850.553	−0.971442	−0.970609	−0.971442	0.970311	0.963045
0.207	0.329	0.126	−0.426					
0.088	0.142	0.053	−0.183					
0.026	0.040	0.016	−0.052					
−0.213	−0.338	−0.127	0.438					

signs of each specific column of C to get the matrix below C. Thus 14.061 is the sum of the elements of the first column of C when multiplied by the signs of the elements of this column, 3077 is obtained by adding the products of the elements in the first column of C with the signs of the elements in the second column of C, etc. The elements of the rows of this matrix are then multiplied by the corresponding elements of the rows of C and added to get the respective values of vCu, which are placed in a matrix at the right of this second matrix. These values are checked by the matrix at the right of C which is obtained with the use of row sums rather than column sums. These four matrices are bordered with the sums of the absolute values of the elements in the rows and columns of C, and with the sum of the absolute values of c_{ij} which is presented in the lower right. This portion of the table corresponds to a part of Table II.

The values of the extreme elements of $C\theta C$ for each value of e_{ij} are then computed by multiplication of the diagonal term of the row sum matrix by the diagonal term of the column sum matrix, or what amounts to the same thing, the values of $\sum_i |c_{ri}| \sum_j |c_{js}|$ are computed and agree with those of Table II. The values of $1 \pm \eta vCu$ are then computed and are placed in a matrix at the right of the $C\theta C$ matrix. A transposition of the (vCu) matrix is effected so that the elements of the recorded matrix correspond to the elements of the $C\theta C$ matrix.

The actual values of the extreme discrepancy matrix are then computed. A negative sign is assigned to certain elements since, in this case, vCu is negative and the element of the discrepancy matrix must be negative in order to attain this extreme value.

In this problem, since η is small, these results are not much better than those of Table II, and they involved considerable additional calculation. In situations such as this, the method of Table II is to be recommended. In other situations, particularly where η is large and when the values of vCu differ appreciably in absolute value from $\sum_i \sum_j |c_{ij}|$, the method just described usually gives much better results.

XII. EXTREME DISCREPANCIES FOR THE CASE OF UNEQUAL BOUNDS

The methods of the last section are useful in treating the case of common bounds but they are not all so useful in handling the general case of unequal bounds. In particular, the method of adjusted matrices indicated in the last section is not usually applicable since the E matrix cannot usually be written in the form uCv. Though the matrix of extreme errors

$$E = \begin{bmatrix} 0.0001 & 0.0002 & 0.0003 \\ 0.0002 & 0.0004 & 0.0006 \\ 0.0003 & 0.0006 & 0.0009 \end{bmatrix}$$

can be written as

$$E = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [0.0001][1 \ 2 \ 3],$$

the matrix of errors of Table I cannot be so written. What can we do in this case?

One answer to this question is to replace each η_{ij} by the largest $\eta_{ij}=\eta$ and then use the methods of the last section. We might take $\eta=0.0015$ rather than 0.0005 in Tables IV, V, VI and VII and thus obtain answers to the problem of Table I. For example, we get from Table V

$$d_{44} = \frac{(0.0015)(850.553)}{1 - (0.0015)(59.378)} = 1.401.$$

These results give bounds, rather than extreme values, and we do not know how much these bounds might be improved. We are still faced with the problem of finding the extreme discrepancies.

Perhaps the simplest method, at least in theory, to use is the method of recalculation. The signs of the error terms are determined as in the last section. Using the η_{ij} of Table I we find that the inversion of

$$A + E = \begin{bmatrix} 0.4940 & 0.2745 & 0.0975 & 0.4600 \\ 0.4350 & 0.4480 & 0.2635 & 0.4110 \\ 0.3940 & 0.5560 & 1.3175 & 0.3860 \\ 0.4240 & 0.4370 & 0.4680 & 0.4680 \end{bmatrix}$$

results in

$(A+E)_{44}^{-1}=15.3156$ so that subtraction of $C_{44}=14.896$ gives $d_{44}=0.419$. This is considerably smaller than the value, 1.401, obtained with the method of common bounds.

This method of recalculation may be considered impracticable by some, and some alternative method may be preferred. However, it should be mentioned here that this method of recalculation may become preferable to other methods as improved techniques of calculation are made available and as machinery is developed for computing inverse matrices quickly.

Alternatively we may use the method of expansion in series. The computation is somewhat more complicated here, however, for the size of the bound of each element must be taken into consideration at each step. Since the values of the errors are usually of the same order, we find it convenient to use a measure of this order as a scalar multiplier. Thus the matrix of errors of Table I, with the signs attached so as to obtain the extreme positive value of d_{44}, can be written

$$E = -\begin{bmatrix} 2 & 1 & 1 & -2 \\ 2 & 2 & 1 & -2 \\ 2 & 2 & 3 & -2 \\ -2 & -2 & -2 & 2 \end{bmatrix} [0.0005] \tag{12.1}$$

or

$$E = H\eta. \tag{12.2}$$

where η is the unit of measurement.

The series then becomes

$$D = \eta CHC + \eta^2 CHCHC + \eta^3 CHCHCHC \cdots, \tag{12.3}$$

and the expansion is carried out as in Table IV.

The method of adjusted matrices, as described above, is not satisfactory in this case since in general the extreme error matrix cannot be written as the product of two vectors.

XIII. MAXIMUM EFFECTS OF COMPUTATIONAL ERRORS

Many methods of computation do not lead to exact values of $A^{-1} = C$, or even to values which can be guaranteed to a fixed number of decimal places, but rather to values which are admittedly approximations to C and which, it is hoped, may be used in place of C. If the approximate matrix is denoted by C_0, what discrepancies between C_0 and C are permissible?

In a paper mentioned above, [26], one of the authors used the method of submatrices in arriving at a five decimal place approximation to C. This approximation, rounded to three decimal places, is

$$C = \begin{bmatrix} 6.485 & 5.152 & 3.311 & -13.625 & 28.573 \\ -4.798 & 6.485 & -0.841 & -\ 0.285 & 12.409 \\ 0.694 & -1.627 & 1.084 & -\ 9.147 & 3.552 \\ -2.086 & -9.095 & -3.299 & 14.896 & 29.376 \\ 14.063 & 22.359 & 8.535 & 28.953 & 73.910 \end{bmatrix} \tag{13.1}$$

(The sums of the absolute values of the row and column elements are placed in the margins for later use.) Is C_0 close enough to C so that it can be used in answering questions of the type discussed in the earlier sections of this paper? Can a bound for the computational errors be found? Can we find a matrix of computational errors which, when added to C_0, gives C to three decimal places?

The question of the accumulation of rounding-off errors has received extensive attention in the literature of the last decade. A complete survey of this material is hardly appropriate to this article but a few of the techniques should be mentioned. Hotelling [11] proposed that an approximation C_0 might serve as the basis of an iterative method which continues until the result is correct to the desired number of places. Satterthwaite [20] has proposed a type of error control in which the norm is reduced to a very small amount by matrix multiplication; von Neumann and Goldstine [17] have studied the accumulation of computational errors in what we call the method of single division [5, p. 99]; Turing [24] studied the accumulation of errors of several common reduction methods using the maximum coefficient. Hotelling [10] provided formulas using norms which give bounds to the difference between C and C_0 and between C and C_m where C_m is the improved approximation to A^{-1} resulting from m iterations. Following the style of Hotelling, but not his precise notation, we write the matrix of computational errors as

$$\Delta = C - C_0 = C(I - AC_0) = CG. \tag{13.2}$$

Using norms we get

$$N(\Delta) \leqq N(C)N(G). \tag{13.3}$$

Unfortunately we do not know $N(C)$ but we can develop an upper bound for it if $N(G)<1$. Since

$$C = C_0[AC_0]^{-1} = C_0[I - G]^{-1} = C_0[I + G + G^2 + \cdots], \tag{13.4}$$

we have

$$\Delta = C - C_0 = C_0[G + G^2 + G^3 + \cdots] \tag{13.5}$$

so

$$N(C) = \frac{N(C_0)}{1 - k} \tag{13.6}$$

$$N(\Delta) = \frac{N(C_0)k}{1 - k}, \tag{13.7}$$

where $k=N(G)$.

This formula is given by Hotelling [10, p. 281]. A similar formula using $M(\Delta)$ is given by Turing [24]. In our notation his formula is

$$M(\Delta) = \frac{nM(C_0)M(G)}{1 - \frac{1}{2}nM(G)}. \tag{13.8}$$

We apply (13.7) and (13.8) to the problem above. We compute $G=I-AC_0$ and get

$$10^6 \cdot G = \begin{bmatrix} -1163 & 436 & 339 & -108 \\ -1114 & 441 & 427 & -87 \\ -1365 & 403 & 911 & -447 \\ -1163 & 418 & 528 & -148 \end{bmatrix} \begin{matrix} 2046 \\ 2069 \\ 3126 \\ 2257 \end{matrix} \tag{13.9}$$

with

$$N_h(C_0) = 25.598, \quad N_b(C_0) = 29.376, \quad M(C_0) = 14.896$$
$$N_h(G) = 0.002860, \quad N_b(G) = 0.003126, \quad M(G) = 0.001365 \tag{13.10}$$

so

$$N_h(\Delta) \leqq 0.073, \; N_b(\Delta) \leqq 0.092, \text{ and } M(\Delta) \leqq 0.082, \tag{13.11}$$

and it follows that no element of C_0 differs from the corresponding element of C by more than 0.073. This is less than many of the extreme errors attributable to inaccuracies of the original data (see Table VII) when $\eta=0.0005$ so there may be little point in making closer approximation to A^{-1} unless it is possible to state the original matrix more accurately.

If the approximation, C_0, were not so good, i.e., if $N(\Delta)$ were considerably larger, it might be wise to obtain a closer approximation to C. Hotelling suggests

$$C_1 = C_0[2I - AC_0] = C_0 + C_0G \tag{13.12}$$

as the next approximation. Then

$$\Delta_1 = C - C_1 = C - C_0 - C_0G = \Delta - C_0G \tag{13.13}$$

and the difference between the approximations is C_0G.

This process can be extended. If it is extended through m iterations we can arrive at Hotelling's formula

$$N(\Delta_m) = \frac{N(C_0)k^{2^m}}{1-k}. \tag{13.14}$$

Hotelling points out that the 2^m exponential of k means that the number of sure decimal places is doubled with each iteration.

Sometimes we wish to provide the Δ which, when added to C_0, gives the value of C to a specified number of decimal places. Though this

can be done by iterative methods, such as the method proposed by Hotelling, it can also be accomplished with the use of (13.5) if C_0 is close enough to C so that the series converges. The method is effective when C_0 is a good approximation to C, as in this case one needs to compute but a few terms of the series to obtain the correct values. This situation usually exists when direct methods or build-up methods are used in obtaining C_0.

In many cases, a single term of the series will be sufficient. This is true in the illustration above since C_0G^2, C_0G^3, $\cdots$ are too small to make any contribution to the third decimal place. It follows that

$$\Delta = C_0G = \begin{bmatrix} -0.002 & 0.001 & 0.000 & -0.001 \\ 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 \end{bmatrix}. \tag{13.15}$$

Addition of Δ to C_0 gives the value of C correct to three decimals.

The formula for (13.5) can be used in another way. Following the type of argument of Section 5 for inherent errors, we let $g_1, g_2, \cdots, g_n$ be the maximum absolute elements in columns $1, 2, \cdots, n$ of G. Then

$$B(G) = \begin{bmatrix} g_1 & g_2 \cdots g_n \\ g_1 & g_2 \cdots g_n \\ \cdots & \cdots \\ g_1 & g_2 \cdots g_n \end{bmatrix}. \tag{13.16}$$

Each row of $B(G^2)$ is $g_1 \sum g, g_2 \sum g, \cdots, g_n \sum g$. Each row of $B(G^p)$ is $g_1(\sum g)^{p-1}, g_2(\sum g)^{p-1}, \cdots, g_n(\sum g)^{p-1}$.

The matrix $\sum_{p=1}^{\infty} B(G)^p)$ is composed of the rows

$$\frac{g_1}{1-\sum g}, \quad \frac{g_2}{1-\sum g}, \cdots, \frac{g_n}{1-\sum g}$$

so that

$$\text{extreme } \Delta_{rs} = \sum_k |C_{0:rk}| \frac{g_s}{1-\sum g}. \tag{13.17}$$

Application of (13.17) using (13.1) and (13.9) yields the result

$$B(\Delta) = \begin{bmatrix} 0.039 & 0.013 & 0.026 & 0.013 \\ 0.017 & 0.005 & 0.011 & 0.006 \\ 0.005 & 0.002 & 0.003 & 0.002 \\ 0.040 & 0.013 & 0.027 & 0.013 \end{bmatrix} \tag{13.18}$$

which, while not as good as (13.15), is much better than the results of (13.11).

REFERENCES

[1] Bartlett, M. S., "An inverse matrix adjustment arising in discriminant analysis," *Annals of Mathematica' Statistics*, 22 (1951), 107–11.

[2] Bowker, Albert H., "On the norm of a matrix," *Annals of Mathematical Statistics*, 18 (1947), 285–88.

[3] Brock, J. E., "Variation of coefficients of simultaneous linear equations," to be published in *Quarterly of Applied Mathematics*.

[4] Dwyer, Paul S., *Errors of Matrix Computations*, presented to Symposium of Solution of Equations, Institute for Numerical Analysis, Los Angeles, California, August 23, 1951.

[5] Dwyer, Paul S., *Linear Computations*, New York, John Wiley and Sons, 1951.

[6] Dwyer, P. S., and MacPhail, M. S., "Symbolic Matrix Derivatives," *Annals of Mathematical Statistics*, 19 (1948), 517–34.

[7] Frazer, R. A., Duncan, W. J., and Collar, A. R., *Elementary Matrices*, Cambridge, Macmillan, 1947.

[8] Frobenius, G., "Über Matrizen aus positiven Elementen I, II," *Sitzungsberichte der Berliner Akademie der Wissenschaften*, (1908), 471–76; (1909) 514–18.

[9] Holley, J. L., "Note on the inversion of the Leontief Matrix," *Econometrica*, 19 (1951), 317–20.

[10] Hotelling, Harold, "Practical problems of matrix calculation," *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1949, 275–93.

[11] Hotelling, Harold, "Some new methods in matrix calculation," *Annals of Mathematical Statistics*, 14 (1943), 1–33.

[12] Janet, M., "Sur les systèmes d'équations aux dérivées partielles," *Journal des Mathématiques*, Series 8, Volume 3 (1920), 65–151.

[13] Lonseth, A. T., "The propagation of error in linear problems," *Transactions of the American Mathematical Society*, 62 (1947), 193–212.

[14] Michal, A. D., *Matrix and Tensor Calculus*, New York, John Wiley and Sons, 1947.

[15] Morgenstern, O., and Woodbury, M. A., "The stability of inverses of input-output matrices," *Econometrica*, 18 (1950), 191–92.

[16] Morrison, I. F., "Variation of coefficients of simultaneous linear equations," *Transactions ASCE*, 113 (1948), 1379.

[17] von Neumann, J., and Goldstine, H. H., "Numerical inverting of matrices of high order," *Bulletin of the American Mathematical Society*, 53 (1947), 1921–99.

[18] Ostrowski, A., "Sur la détermination des bornes inférieures pour une classe des déterminants," *Bulletin des Sciences Mathématiques*, II, 61 (1937), 19–32.

[19] Ostrowski, A., "Über die Determinanten mit überwiegender Hauptdiagonale," *Comm. Math. Helv.*, 10 (1937), 69–96.

[20] Satterthwaite, F. E., "Error control in matrix calculation," *Annals of Mathematical Statistics*, 15 (1944), 373–87.

[21] Sherman, J., and Morrison, W. J., "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix," *Annals of Mathematical Statistics*, 21 (1950), 124–25.

[22] Sherman, J., and Morrison, W. J., "Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix," (Abstract), *Annals of Mathematical Statistics*, 20 (1949), 621.

[23] Taussky, O., "A recurring theorem on determinants," *American Mathematical Monthly*, 56 (1949), 672–76.

[24] Turing, A. M., "Rounding off errors in matrix processes," *Quarterly Journal of Mechanics and Applied Mathematics*, 1 (1948), 287–308.

[25] Waugh, F. V., "Inversion of the Leontief matrix by power series," *Econometrica*, 18 (1950), 142–54.

[26] Waugh, F. V., "The minimum cost dairy feed," *Journal of Farm Economics*, 33 (1951), 299–310.

[27] Waugh, F. V., and Dwyer, P. S., "Compact computation of the inverse of a matrix," *Annals of Mathematical Statistics*, 16 (1945), 259–71.

[28] Weiner, B. L., "Variations of coefficients of simultaneous linear equations," *Transactions ASCE*, 113 (1948), 1349.

[29] Wittmeyer, H., "Einfluss der Änderung einer Matrix, etc.," *Zeitschrift für angewandte Mathematik und Mechanik.*, 16 (1936), 287–300.

[30] Woodbury, M. A., "Inverting modified matrices," Statistical Research Group, Princeton University, *Memorandum Report* 42 (1950).

[31] Woodbury, M. A., "Properties of Leontief-type input-output matrices," Economics Research Project, Princeton University.

A MULTIPLE GROUP LEAST SQUARES' PROBLEM AND THE SIGNIFICANCE OF THE ASSOCIATED ORTHOGONAL POLYNOMIALS

BRADFORD F. KIMBALL
New York State Department of Public Service

A multiple group least squares' problem is presented and solved which involves determination of certain coefficients separately for each of several samples with a final coefficient given a single optimum value for all samples grouped together.

The computation of orthogonal polynomials implicit in the normal equations is discussed relative to the forward solution of these equations. Certain significant advantages of computing orthogonal predictors are pointed out.

A numerical example of the computation of the orthogonal polynomials in the case of a weighted fit is presented with application to determination of the confidence band of a fitted trend curve.

I. A MULTIPLE GROUP LEAST SQUARES' PROBLEM

SUPPOSE a cost-distance relationship is sought in a study of motor carrier costs. Cost data from samples of several companies are available but they are from different regions and apply to different weight ranges of shipments. A quadratic polynomial distance function is to be fitted by least squares, where the distance variable x is taken as the square root of the distance, and the dependent cost variable is cost per 100 lbs.

There is reason to take the coefficient of the squared term as constant over all samples, while the linear coefficients are to be fitted to the separate samples. How does one solve the problem?

The answer is furnished by the following analysis: In the interests of simplicity limit the problem to three samples. One then seeks three cost-distance functions of the type

$$\begin{aligned} u^{(1)} &= \alpha^{(1)} + \beta^{(1)}x + \gamma x^2 \\ u^{(2)} &= \alpha^{(2)} + \beta^{(2)}x + \gamma x^2 \\ u^{(3)} &= \alpha^{(3)} + \beta^{(3)}x + \gamma x^2 \end{aligned} \tag{1.1}$$

which give a best fit to the data in the least squares' sense.

The costs occur as three series $y_j^{(s)}$, $s=1, 2, 3$ with $N^{(s)}$ values in the sth series. The moments are designated by

$$a_{ik}^{(s)} = S[z_{ij}^{(s)} z_{kj}^{(s)}], \qquad a_{i0}^{(s)} = S[z_{ij}^{(s)} y_j^{(s)}] \tag{1.2}$$

where

$$z_{1j}^{(s)} = 1, \qquad z_{2j}^{(s)} = x_j^{(s)}, \qquad z_{3j}^{(s)} = (x_j^{(s)})^2.$$

Thus the matrix of the normal equations for each sample considered separately will be

$$\|a_{i1}^{(s)} \quad a_{i2}^{(s)} \quad a_{i3}^{(s)}, \quad a_{i0}^{(s)}\|, \qquad i = 1, 2, 3. \tag{1.3}$$

For each sample we first neglect the squared term in x and obtain solutions

$$U^{(s)} = B_1^{(s)} + B_2^{(s)}x \tag{1.4}$$

of the normal system

$$\|a_{i1}^{(s)} \quad a_{i2}^{(s)}, \quad a_{i0}^{(s)}\|, \qquad i = 1, 2. \tag{1.5}$$

As the next step determine a pair of solutions $A_1^{(s)}$, $A_2^{(s)}$ of the system

$$\|a_{i1}^{(s)} \quad a_{i2}^{(s)}, \quad a_{i3}^{(s)}\|, \qquad i = 1, 2 \tag{1.6}$$

where the moments of the dependent variable on the right have been replaced by the moments involving z_3.

It will be found that the first two equations of the system (1.3) are satisfied by the three coefficients

$$B_1^{(s)} - tA_1^{(s)}, \quad B_2^{(s)} - tA_2^{(s)}, \quad t$$

for *any* value of t. Thus the two following equations

$$\begin{aligned} a_{11}^{(s)}(B_1^{(s)} - tA_1^{(s)}) + a_{12}^{(s)}(B_2^{(s)} - tA_2^{(s)}) + a_{13}^{(s)}t &= a_{10}^{(s)} \\ a_{21}^{(s)}(B_1^{(s)} - tA_1^{(s)}) + a_{22}^{(s)}(B_2^{(s)} - tA_2^{(s)}) + a_{23}^{(s)}t &= a_{20}^{(s)} \end{aligned} \tag{1.7}$$

will be satisfied for *each* sample, for *any* value of t. Then the quadratic curve, for any constant value of t,

$$u(t) = B_1^{(s)} - tA_1^{(s)} + (B_2^{(s)} - tA_2^{(s)})x + tx^2$$

will have linear terms which will be the optimum terms in the least squares' sense for each s for the sample in question. This means that

$$\partial S^{(s)}/\partial(B_1^{(s)} - tA_1^{(s)}) = \partial S^{(s)}/\partial(B_2^{(s)} - tA_2^{(s)}) = 0 \tag{1.8}$$

where $S^{(s)}$ denotes the sum of the squares of the residuals for the sth sample.

Thus our problem will be solved if t can be chosen to give an optimum value for all the samples. This optimum is taken as the value

which will minimize the sum of the squares of the residuals over *all* the samples lumped together.

Now for a single sample the total derivative of the sum of the squares of the residuals with respect to t can be shown to be

$$dS^{(s)}/dt = -2(Y_3^{(s)} - q_3^{(s)}t) \tag{1.9}$$

where

$$\begin{aligned} q_3^{(s)} &= a_{33}^{(s)} - A_1^{(s)}a_{31}^{(s)} - A_2^{(s)}a_{32}^{(s)} \\ Y_3^{(s)} &= a_{30}^{(s)} - B_1^{(s)}a_{31}^{(s)} - B_2^{(s)}a_{32}^{(s)}. \end{aligned} \tag{1.10}$$

Using (1.8) this is proved as follows: Omitting the superscripts, write

$$u(t) = P(t) + xQ(t) + tx^2,$$
$$P(t) = B_1 - tA_1, \qquad Q(t) = B_2 - tA_2$$

and the total derivative of S as to t is given by

$$\frac{dS}{dt} = \frac{\partial S}{\partial P}\frac{dP}{dt} + \frac{\partial S}{\partial Q}\frac{dQ}{dt} + \frac{\partial S}{\partial t}$$

where the last partial derivative involves t only as coefficient of x^2 in $u(t)$. From (1.8)

$$\partial S/\partial P = \partial S/\partial Q = 0.$$

Hence

$$dS/dt = \partial S/\partial t = 2\sum (u - y)x^2 = -2(\sum x^2y - \sum x^2u)$$

which easily reduces to (1.9).

Now note that (cf. (1.7))

$$\begin{aligned} &-(Y_3^{(s)} - q_3^{(s)}t) \\ &= a_{31}^{(s)}(B_1^{(s)} - tA_1^{(s)}) + a_{32}^{(s)}(B_2^{(s)} - tA_2^{(s)}) + (a_{33}^{(s)}t - a_{30}^{(s)}). \end{aligned} \tag{1.11}$$

Using (1.9) the derivative of the sum of the squares of the residuals taken over all samples is given by summing the expression on the right over all samples with t constant. Setting this equal to zero

$$\hat{t} = \gamma = \frac{Y_3^{(1)} + Y_2^{(2)} + Y_3^{(3)}}{q_3^{(1)} + q_3^{(2)} + q_3^{(3)}}. \tag{1.12}$$

Hence *the required coefficients of the quadratics (1.1) are given by*

$$\alpha^{(s)} = B_1^{(s)} - \gamma A_1^{(s)}, \qquad \beta^{(s)} = B_2^{(s)} - \gamma A_2^{(s)} \tag{1.13}$$

where γ is determined by (1.12).

II. ANALOGY WITH METHOD OF ORTHOGONAL POLYNOMIALS

Light is thrown on the problem of finding the sampling variance of γ by bringing out the orthogonal relations inherent in the method. The method of "orthogonal polynomials" would proceed by first determining orthogonal functions ϕ_i of the variables z_i (see (1.2) above). The requirement is that

$$S\phi_1\phi_2 = S\phi_1\phi_3 = S\phi_2\phi_3 = 0 \tag{2.1}$$

where the summations are taken over the N values of the independent variables z_{ij} which correspond to the data series on y_j in a particular sample. The author has shown [6, Sec. 3] that a set of such orthogonal functions may be taken as

$$\phi_1 = z_1, \qquad \phi_2 = z_2 - Mz_1, \qquad \phi_3 = z_3 - A_1z_1 - A_2z_2 \tag{2.2}$$

where

$$M = a_{12}/a_{11} \tag{2.3}$$

and A_1 and A_2 are defined by (1.6) above, for a single sample.

The solution is sought in the form

$$u = t_1\phi_1 + t_2\phi_2 + t_3\phi_3 \tag{2.4}$$

and the normal equations for a single sample reduce to

$$t_1S\phi_1^2 = S\phi_1y, \qquad t_2S\phi_2^2 = S\phi_2y, \qquad t_3S\phi_3^2 = S\phi_3y. \tag{2.5}$$

An examination of the summation products $S\phi_3^2$ and $S\phi_3y$ shows that for a single sample

$$S\phi_3^2 = q_3, \qquad S\phi_3y = Y_3 \tag{2.6}$$

where q_3 and Y_3 are given by (1.10). This is proved in the general case in [6], formulas (4.5), (4.6) and discussion of (3.2).

Hence if one denotes by $\phi_i^{(s)}$ the orthogonal functions applicable to the sth sample, one can write

$$\gamma = \frac{S\phi_3^{(1)}y^{(1)} + S\phi_3^{(2)}y^{(2)} + S\phi_3^{(3)}y^{(3)}}{q_3^{(1)} + q_3^{(2)} + q_3^{(3)}} \tag{2.7}$$

where the denominator involves only the independent variables z_i.

Thus one can proceed to find the sampling variance of γ as follows: The first order variation in γ due to variations in y_j is given by

$$(2.8) \qquad \delta\gamma = [S\phi_3^{(1)}\delta y^{(1)} + S\phi_3^{(2)}\delta y^{(2)} + S\phi_3^{(3)}\delta y^{(3)}]/Q$$

where for brevity Q denotes the denominator of (2.7). Making the usual assumption that the y_j are independent with constant variance,

$$E[(\delta\gamma)^2] = E[(\delta y)^2][S(\phi_3^{(1)})^2 + S(\phi_3^{(2)})^2 + S(\phi_3^{(3)})^2]/Q^2$$

which by (2.6) reduces to

$$E[(\delta\gamma)^2] = E[(\delta y)^2]/Q.$$

Hence *the estimated sampling variance of* γ *is given by*

$$(2.9) \qquad \text{Est. Variance of } \gamma = V/(q_3^{(1)} + q_3^{(2)} + q_3^{(3)})$$

where V *equals the sum of the squares of the residuals divided by* $(N^{(1)}+N^{(2)}+N^{(3)}-7)$. This last denominator represents the proper number of degrees of freedom since it is the total number of elements in the three samples (see (1.1) above) decreased by the total number of coefficients determined in the fitting process.

The above method can be applied equally well if different statistical weights are assigned to the observations on the dependent variable. The only difference in procedure would be that the moments would be *weighted* moments, and that one would deal with a *weighted* sum of the squares of the residuals. The formulas (2.1), (2.5) and (2.6) would involve *weighted* sums of the orthogonal functions, but the determination of the explicit form of the orthogonal polynomials (2.2) would remain unchanged since the weighting process would be included in the primary moments which determine M, A_1 and A_2.

Problems of the type discussed in Sections I and II might very well occur in many other fields of application of least squares' theory. The essential characteristics of the problem are: A determination of a set of coefficients of fitted functions so that they separately give optimum fit to each particular sample; a determination of another coefficient so that it gives optimum fit to the group of samples lumped together.

The procedure can be further generalized to apply to two or more sets of coefficients, where there is more than one coefficient in the second set, and to cases where the grouping of samples is performed more than once. This generalization involves some changes in procedure and may be treated by the author in a later paper.

III. SOLUTION BY MEANS OF THE DOOLITTLE ALGORITHM

The Doolittle algorithm supplemented by columns used for computing the elements of the reciprocal matrix, furnishes an efficient means of computing the necessary quantities in the preceding problem.

We use a slight modification of Dwyer's notation and first write the moment matrix supplemented by the array for computing elements of the reciprocal matrix, for three variables. This is termed the *reference matrix* in Schedule I below. This is followed by the *computation algorithm* where for brevity Dwyer's "secondary" subscripts have been omitted from the b's and the g's, and a zero subscript has been used to denote the dependent variable [2, pp. 99–118 and 170].

SCHEDULE I

Reference Matrix

Row	y	z_1	z_2	z_3	z_1	z_2	z_3
I	a_{10}	a_{11}	a_{12}	a_{13}	1	0	0
II	a_{20}		a_{22}	a_{23}	0	1	0
III	a_{30}			a_{33}	0	0	1

Computation Algorithm

$S(\mathrm{I}) = \mathrm{I}$	a_{10}	a_{11}	a_{12}	a_{13}	1	0	0
I′	b_{10}	1	b_{12}	b_{13}	—	—	—
$S(\mathrm{II})$	g_{20}		g_{22}	g_{23}	$-M$	1	0
II′	b_{20}		1	b_{23}	—	—	—
$S(\mathrm{III})$	g_{30}			g_{33}	$-A_1$	$-A_2$	1
III′	b_{30}			1	—	—	—

For the purpose of determining the solution in terms of orthogonal polynomials it is not necessary to carry the primed rows through the auxiliary part of the computational schedule. The orthogonal polynomials are given explicitly by the elements in the summation rows S(I), S(II) and S(III). Thus

$$\phi_1 = z_1, \qquad \phi_2 = z_2 - Mz_1, \qquad \phi_3 = z_3 - A_1z_1 - A_2z_2. \tag{3.1}$$

The multipliers which apply to these orthogonal functions are the b's of the first column. Thus the fitted function is

$$u = b_{10}\phi_1 + b_{20}\phi_2 + b_{30}\phi_3 \tag{3.2}$$

and for

$$u = h_1^{(3)}z_1 + h_2^{(3)}z_2 + h_3^{(3)}z_3 \tag{3.3}$$

the forward solution for the three variable fit is

$$\begin{aligned} h_1^{(3)} &= b_{10} - Mb_{20} - A_1b_{30} \\ h_2^{(3)} &= \quad b_{20} - A_2b_{30} \\ h_3^{(3)} &= \quad b_{30}. \end{aligned} \tag{3.4}$$

For two variables omit the third column

It has been shown [6, formula (4.5) and discussion of (3.2)] that the g elements in the summation rows (not including the auxiliary schedule) are related to the orthogonal functions as follows:

$$(3.5) \qquad g_{r0} = Sy\phi_r, \qquad g_{rk} = Sx_k\phi_r, \qquad g_{rr} = S\phi_r^2$$

where summations are taken over range of observations fitted.

Thus from (2.6)

$$(3.6) \qquad q_3 = S\phi_3^2 = g_{33}, \qquad Y_3 = Sy\phi_3 = g_{30}.$$

Hence the solution sought, in terms of the notation of the Doolittle algorithm is

$$(3.7) \qquad \gamma = \frac{g_{30}^{(1)} + g_{30}^{(2)} + g_{30}^{(3)}}{g_{33}^{(1)} + g_{33}^{(2)} + g_{33}^{(3)}}$$

where the g's are determined from separate Doolittle algorithms carried through for each sample, and

$$(3.8) \qquad \begin{aligned} \alpha^{(s)} &= b_{10}^{(s)} - M^{(s)}b_{20}^{(s)} - \gamma A_1^{(s)}, \\ \beta^{(s)} &= \qquad\qquad b_{20}^{(s)} - \gamma A_2^{(s)}, \qquad\qquad s = 1, 2, 3 \end{aligned}$$

give the first two coefficients for each sample, adjusted for γ.

IV. ADVANTAGES OF COMPUTING THE ORTHOGONAL FUNCTIONS

We have seen in the last section that the explicit determination of a set of orthogonal functions inherent in the least squares' solution is possible by supplementing the array of moments used in the reference matrix of the Doolittle algorithm by the identity matrix (with diagonal elements equal to unity and all other elements equal to zero). This process is entirely general and will serve to determine a set of orthogonal functions for any number of independent variables or "predictors."

One may well ask why it is not desirable to complete the computation for the elements of the reciprocal matrix. In some cases this would probably be desirable, for example in a problem of linear correlation where the number of predictors is fixed [4]. However, in many cases one is experimenting with the question of what predictors are suitable, and hence a process by which new variables can be introduced or dropped with a minimum of computation is highly desirable. As we shall presently see, the computation of the coefficients of the orthogonal functions offers an efficient means for such a process of iteration or "build up."

We introduce here a general notation. The general problem is taken as that of fitting a linear function u of N observations on a set of n variables z_i (see (1.2) above) and the matrix of the normal equations is given by (1.3) with no superscripts since we deal only with one sample. The fitted function is written

$$u_n = \beta_1^{(n)} z_1 + \beta_2^{(n)} z_2 + \beta_3^{(n)} z_3 + \cdots + \beta^{(n)} z_n \tag{4.1}$$

where the superscript here denotes the number of predictors used in the fitting process. Thus if only three are used,

$$u_3 = \beta_1^{(3)} z_1 + \beta_2^{(3)} z_2 + \beta_3^{(3)} z_3.$$

The same function expressed in terms of the orthogonal functions is written as

$$u_n = t_1\phi_1 + t_2\phi_2 + t_3\phi_3 + \cdots + t_n\phi_n \tag{4.2}$$

and in this case the solution for three variables is given by using the first three terms of the *same* series.

Denote by $\alpha_i^{(k)}$ the coefficients which determine the orthogonal functions. Thus

$$\begin{aligned}
\phi_1 &= z_1 \\
\phi_2 &= z_2 - \alpha_1^{(1)} z_1 \\
\phi_3 &= z_3 - \alpha_1^{(2)} z_1 - \alpha_2^{(2)} z_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\phi_n &= z_n - \alpha_1^{(n-1)} z_1 - \alpha_2^{(n-1)} z_2 - \cdots - \alpha_n^{(n-1)} z_n.
\end{aligned} \tag{4.3}$$

From the above it follows that

$$\begin{aligned}
\beta_1^{(n)} &= t_1 - t_2\alpha_1^{(1)} - t_3\alpha_1^{(2)} - \cdots - t_n\alpha_1^{(n-1)} \\
\beta_2^{(n)} &= \quad t_2 \quad - t_3\alpha_2^{(2)} - \cdots - t_n\alpha_2^{(n-1)} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\beta_{n-1}^{(n)} &= \qquad\qquad\qquad\qquad\quad t_{n-1} - t_n\alpha_{n-1}^{(n-1)} \\
\beta_n^{(n)} &= \qquad\qquad\qquad\qquad\qquad\quad t_n
\end{aligned} \tag{4.4}$$

where the multipliers t_i are the same as the b elements taken from the first column of Schedule I. Using the abbreviated Dwyer notation,

$$t_1 = b_{10},\ t_2 = b_{20},\ \cdots,\ t_n = b_{n0}. \tag{4.5}$$

We have seen in the last section that the coefficients of the orthogonal functions for the first three variables are given by the elements in the "summation" rows of the auxiliary schedule of the Doolittle algo-

rithm. This is known to be true in the general case [6, Sec. 3]. Hence in terms of our general notation the elements in the summation rows of this auxiliary schedule appear as

SCHEDULE II

Auxiliary to Doolittle Algorithm

	Row	z_1	z_2	z_3	...	z_n
	S(I)	1	0	0	...	0
	S(II)	$-\alpha_1^{(1)}$	1	0	...	0
(4.6)	S(III)	$-\alpha_1^{(2)}$	$-\alpha_2^{(2)}$	1	...	0
	. . . . .					
	$S(n)$	$-\alpha_1^{(n-1)}$	$-\alpha_2^{(n-1)}$	$-\alpha_3^{(n-1)}$	...	1

The relationships (4.4) to (4.6) offer precisely the "forward solution" suggested by Dwyer [2, p. 170]. In other words, if one uses only two columns in the array (4.4) the solution of the two variable problem is obtained. Addition of the third column gives the three variable solution, and so forth.

If one uses the fact that the t's are the multipliers of the orthogonal functions in the solution (4.2), one realizes that the t's are statistically independent in the sense of having zero covariance over the range used in fitting. Assuming no errors in the predictors, the sampling variance weight of t_k is known to be the reciprocal of the antecedent g_{kk} (see Schedule I). Hence the variance weight of any coefficient $\beta_k^{(n)}$ of the n variable problem is given by

$$\text{Variance Weight of } \beta_k^{(n)} = 1/g_{kk} + (\alpha_k^{(k)})^2/g_{k+1,k+1} + \cdots + (\alpha_k^{(n-1)})^2/g_{nn}. \tag{4.7}$$

The advantage of this formula is that to find the variance weight of the coefficient $\beta_k^{(n-1)}$ of the same predictor, when another prediction variable is added, one merely adds the term

$$(\alpha_k^{(n)})^2/g_{n+1,n+1}.$$

A similar formula can be set up by the researcher for the covariance of two coefficients.

As more predictors are added, the variances of all coefficients $\beta_i^{(n)}$ are tabulated and partial correlations can easily be found from the following relation. Denote by $S^{(n)}$ the sum of the squares of the residuals after fitting a linear function of n predictors. Let $S_k^{(n-1)}$ denote the sum of the squares of the residuals when the kth predictor z_k is omitted in the fitting. The following equation then holds [1, p. 163].

$$S_k^{(n-1)} = S^{(n)} + (\beta_k^{(n)})^2 \text{ divided by Variance Wt. of } \beta_n^{(n)}. \tag{4.8}$$

From this the partial correlation of y relative to z_k for the n variable solution can be computed from the relation [3, p. 180]

$$1 - r^2_{0k \cdot 12 \cdots n} = [S^{(n)}/S_k^{(n-1)}](N - n - 1)/(N - n). \tag{4.9}$$

The above analysis does not explicitly involve the orthogonal functions and thus does not involve new theory. The results are obtained simply from a study of the relations (4.4) above and constitute an argument for the computation of Schedule II.

An explicit result following from the fact that the computations summarized in Schedule II give the orthogonal functions is that in the case of the linear problem the orthogonal functions furnish a new set of predictors which are linear transforms of the original variables, and possess zero covariance over the range of observed values used in the fitting process. In this connection it should of course be recognized that these orthogonal functions depend upon the *order* in which the original variables have been considered. For example, if z_k were taken as the first variable in the fitting process, we should have $\phi_1 = z_k$. Hence any change in the order in which the predictors are considered, means a change in the corresponding set of orthogonal functions computed by the above method.

Knowledge of the explicit nature of these "orthogonal predictors" may throw light upon the nature of the original variables used. Indeed, in researches in psychology and economics it has been found that the determination of such "principal components" offers a starting point for simplifying inherent structural relationships.

Analysis by means of orthogonal functions and their relation to the Doolittle algorithm makes it much easier to generalize the type of problem discussed in Sections 1–3 and offers a solution which can be carried out in a systematic manner.

Perhaps the most obvious application is to problems of fitting of polynomial curves which are to be used for purposes of forecasting. In such cases it is often desirable to estimate the variance of the fitted polynomial at several points on the curve for the purpose of determining a sampling confidence band.

For curves of any complexity, say of the third or fourth degree this is simpler to do in terms of the orthogonal functions implicitly used in the fitting process. Recalling that the reciprocals of the g_{kk} elements in Schedule I constitute the variance weights of the multipliers t_i of the orthogonal functions, which are now orthogonal polynomials of a single variable, say x; the estimated variance of the fitted polynomial u at any point x_i is approximated by (cf. [5], formula (21))

$$s_u^2 = V_n[\phi_1^2/a_{11} + \phi_2^2/g_{22} + \cdots + \phi_n^2/g_{nn}] \tag{4.10}$$

where

$$V_n = S^{(n)}/(N - n)$$

applicable to a polynomial of $(n-1)$st degree.

With the orthogonal polynomials determined by Schedule II, where now

$$z_1 = 1, z_2 = x, z_3 = x^2, \cdots, z_n = x^{n-1},$$

their values at a series of points can be systematically computed with considerable ease, and the resulting numerical values are then squared. There remains merely to divide each one of these n values by the corresponding g_{kk} taken directly from the Doolittle algorithm. The sum gives the variance weight s_u^2/V_n. If the degree of the polynomial is increased by one, the new variance weight involves simply the addition of the term

$$\phi_{n+1}^2/g_{n+1,n+1}$$

to what has already been computed.

The other alternative is to find the covariances and variances of the coefficients $\beta_k^{(n)}$ by completing the computation of the reciprocal matrix and apply them to squares and cross products of the original variables at each point. This is not only a much more complicated process, but also it will have to be repeated from the beginning, if a polynomial of another degree be fitted.

V. NUMERICAL EXAMPLE

As a numerical example we have used a problem presented by Snedecor to illustrate the method of orthogonal polynomials [8]. Data are taken over 11 consecutive days on weight of chick embryos and a succession of polynomial curves to the fourth degree are fitted. We shall use the same data with the exception that we shall suppose that inadvertently no weight was recorded on one of the days (taken as the 11th day which happens to be at the central position). We shall use a weighted fit with statistical weights equal to unity assigned to days when weights were recorded, and a statistical weight of zero assigned to the 11th day when no weight was recorded.

Data to be Fitted

Day:	6	7	8	9	10	11	12	13	14	15	16
y:	29	52	79	125	181	—	425	738	1130	1882	2812
x:	10	9	8	7	6	5	4	3	2	1	0

This procedure enables one to use "factorial moments" in setting up the matrix of the normal equations. We assign values of x such that $x=0$ on the 16th day, at the extreme where y is large, and thus $x=10$ corresponds to the 6th day which is the first day that an observation was recorded. Thus there is no value of y at $x=5$, and statistical weight $w=0$ at $x=5$.

We use Lipps factorial moments [7] which might be termed "reduced" factorial moments or "reduced cumulatives." They represent successive cumulations from $x=10$ towards $x=0$ where each successive cumulation omits the last subtotal in the previous cumulation. The moments of the weighted independent variable x are

$$S_0 = \sum w = 10, \qquad S_1 = \sum w\binom{x}{1} = 50,$$

$$S_2 = \sum w\binom{x}{2} = 155, \text{ etc.}$$

In the case of the dependent variable u the moments are

$$M_0 = \sum wy = 7.453, \qquad M_1 = \sum w\binom{x}{1}y = 11.407,$$

$$M_2 = \sum w\binom{x}{2}y = 16.623, \text{ etc.}$$

Since the major interest is in the plotting of the curves, the matrix of the normal equations is set up to determine successive *differences* of the polynomials at $x=0$. We fit

$$u_6 = \beta_1^{(5)} + \beta_2^{(5)}(x) + \beta_3^{(5)}\binom{x}{2} + \beta_4^{(5)}\binom{x}{3} + \beta_5^{(5)}\binom{x}{4}$$

with

$$z_1 = 1, \quad z_2 = (x), \quad z_3 = \binom{x}{2}, \quad z_4 = \binom{x}{3}, \quad z_5 = \binom{x}{4}.$$

The moments involving the dependent variable in the matrix are then

$$a_{k0} = \sum w \binom{x}{k-1} y = M_{k-1}, \quad k = 1, 2, \cdots, 5.$$

The moments of the independent variable in the matrix are

$$a_{ki} = \sum w \binom{x}{k-1}\binom{x}{i-1}, \qquad k, i = 1, 2, \cdots, 5.$$

Using the formula

$$\binom{x}{r}\binom{x}{s} = \binom{r+s}{r}\binom{x}{r+s} + \binom{s}{1}\binom{r+s-1}{r-1}\binom{x}{r+s-1} + \binom{s}{2}\binom{r+s-2}{r-2}\binom{x}{r+s-2}\cdots$$

with $r \leqq s$, the matrix elements a_{ki} are easily resolved in terms of the reduced factorial moments S_i. For example

$$a_{1i} = a_{i1} = S_{i-1}, \quad a_{34} = 4S_4 + 30S_5 + 60S_6 + 35S_7, \quad \text{etc.}$$

Substitution of the values of the reduced factorial moments S_i yields the following matrix for the normal equations:

y	z_1	z_2	z_3	z_4	z_5
7.453	10	50	155	320	457
11.407	—	360	1270	2788	4133
16.623	—	—	4817	11054	16912
22.705	—	—	—	26234	41233
25.687	—	—	—	—	66327

Following the computational procedure outlined in Schedule I, with the identity matrix placed at the right of the normal equation matrix, we record in Schedule III the numerical values found in the "summation" and "primed" rows of the computational algorithm.

The formula for the sum of the squares of the residuals, to be used in the forward solution is (cf. Schedule I and [1, p. 164])

$$S^{(5)} = \sum wy^2 - a_{10}b_{10} - g_{20}b_{20} - \cdots - g_{50}b_{50}$$
$$= 13.509609 - 13.506125 = .003484$$

and the variance of fit is

$$V_5 = .003484/5 = .000697.$$

SCHEDULE III

Computational Algorithm, Five Parameters

	Row	y	From Normal Equation Matrix						From Identity Matrix					
			x_1	x_2	x_3	x_4	x_5		x_1	x_2	x_3	x_4	x_5	
S	(I)=I	7.453	10	50	155	320	457		1	0	0	0	0	0
	I'	0.7453	1	5	15.5	32	45.7		—	—	—	—	—	—
S	(II)	−25.858	—	110	495	1188	1848		−5.0	1	0	0	0	0
	II'	− 0.235073	—	1	4.5	10.8	16.8		—	—	—	—	—	—
S	(III)	17.4625	—	—	187	748	1512.5		7.0	−4.5	1	0	0	0
	III'	0.0338235	—	—	1	4	8.088235		—	—	—	—	—	—
S	(IV)	− 6.3746	—	—	—	171.6	600.6		−6.0	7.2	−4.0	1	0	0
	IV'	− 0.03714802	—	—	—	1	3.5		—	—	—	—	—	—
S	(V)	0.5696	—	—	—	—	60.1446		2.63235	−5.60294	5.911765	−3.5	1	1
	V'	0.00947051	—	—	—	—	1		—	—	—	—	—	—
			(1)	(2)	(2)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)

The forward solution for finding the β's from the relations (4.4) and (4.5) would be facilitated by a separate computational schedule which we do not show here. This would be obtained from Schedule III by multiplying the elements b_{k0} in the first column corresponding to the primed rows by the figures which appear in the preceding "summation" rows obtained from the identity matrix (Cols. 8 to 12). If a long computational sheet is used, this might be placed to the right of Schedule III to avoid copying the multipliers b_{k0}.

One can summarize the numerical results as follows: The values of $\beta_i^{(5)}$ give the differences at $x=0$. They are

$$u_5(0) = 2.822632, \quad \Delta u_5 = -.975823, \quad \Delta^2 u_5 = .2979619$$
$$\Delta^3 u_5 = -.0702948, \quad \Delta^4 u_5 = .00947051.$$

The orthogonal polynomials implicit in the solution defined by the elements in the summation rows of the auxiliary part of Schedule III, are

$$\phi_1 = 1, \quad \phi_2 = x - 5.0, \quad \phi_3 = \binom{x}{2} - (4.5)x + 7$$

$$\phi_4 = \binom{x}{3} - 4\binom{x}{2} + (7.2)x - 6$$

$$\phi_5 = \binom{x}{4} - (3.5)\binom{x}{3} + (5.911765)\binom{x}{2} - (5.60294)x + 2.68235$$

which means that the quartic can be expressed as (see (4.5) above)

$$u_5 = t_1\phi_1 + t_2\phi_2 + t_3\phi_3 + t_4\phi_4 + t_5\phi_5.$$

Subtotals obtained from the forward solutions of the β's give the differences of the intermediate polynomial curves fitted to the data. These are

Straight Line: $u_2 = 1.920665, \quad \Delta u_2 = -.235073,$

Quadratic: $u_3 = 2.574341, \quad \Delta u_3 = -.655294, \quad \Delta^2 u_3 = .0933824,$

Cubic: $u_4 = 2.797229, \quad \Delta u_4 = -.922760, \quad \Delta^2 u_4 = .2419744,$
$\Delta^3 u_4 = -.03714802.$

Using (4.10) and the above computations the first approximation to the variance of the fitted quartic at any point x is given by

$$s_u^2 = (.000697)[0.1 + [\phi_2(x)]^2/110 + [\phi_3(x)]^2/187 + [\phi_4(x)]^2/(171.6) + [\phi_5(x)]^2/(60.1446)].$$

The above set of orthogonal polynomials will of course be valid for any series of y data arranged at 10 equally spaced intervals with no value recorded at the central value of x, where $x=0$ at one extreme and $x=10$ at the other extreme. Because of the simple nature of the distribution of the statistical weights over the scale of x, it may be of interest to record the orthogonal polynomials more in detail. The exact values of the last three coefficients of ϕ_5 are

$$5.911765 = (100.5)/17, \quad 5.60294 = (95.25)/17, \quad 2.68235 = (45.6)/17.$$

Since the weighting is symmetrical about the central value of x, $x=5$, the orthogonal polynomials will satisfy the usual symmetrical relations. They can be written as:

$$\phi_2 = (x-5), \qquad 2\phi_3 = (x-5)^2 - 11,$$

$$6\phi_4 = (x-5)^3 - (17.8)(x-5),$$

$$24\phi_5 = (x-5)^4 - \frac{443}{17}(x-5)^2 + \frac{1544.4}{17}.$$

REFERENCES

[1] Deming, William Edwards, *Statistical Adjustment of Data*, New York, John Wiley and Sons, 1943, 158 and 164.

[2] Dwyer, Paul S., *Linear Computations*, New York, John Wiley and Sons, 1951, 99–118 and 170.

[3] Ezekiel, Mordecai, *Methods of Correlation Analysis*, New York, John Wiley and Sons, 1930, 180.

[4] Hotelling, Harold, "Some new methods in matrix calculation," *Annals of Mathematical Statistics*, 14 (1943), 1–34.

[5] Kimball, B. F., "Orthogonal polynomials applied to least square fitting of weighted observations," *Annals of Mathematical Statistics*, 11 (1940), 348–52.

[6] Kimball, B. F., "Note on computation of orthogonal predictors," to appear in the June 1953 issue of the *Annals of Mathematical Statistics*.

[7] Sasuly, Max, *Trend Analysis of Statistics*, Washington, D. C., The Brookings Institution, 1934, 146.

[8] Snedecor, George W., *Statistical Methods*, Ames, Iowa, Collegiate Press, Inc., 1937, 279–88.

ESTIMATION OF THE PARAMETERS OF TYPE III POPULATIONS FROM TRUNCATED SAMPLES

DES RAJ

University of Lucknow and Indian Statistical Institute, Calcutta

The problem considered is the estimation of the mean and variance of Type III populations from singly and doubly truncated samples with known truncation points when the number of unmeasured observations is (i) unknown for each tail, (ii) known separately for each tail, (iii) known jointly for the two tails. For singly truncated samples, estimation of all three parameters has also been considered. Estimating equations are obtained by the method of moments as well as by the method of maximum likelihood. Previous results for the normal and Type III populations are particular cases of the results obtained in this paper. Numerical examples are given.

THE problem of estimating the mean and variance of normal populations from truncated samples has been considered by Pearson and Lee [10], Lee [9], Fisher [7], Stevens [12], Cochran [1], Hald [8], Cohen [2] and [4], and the author [5]. The corresponding problem for the Type III population, of which the normal population is a particular case, is considered by Cohen [2] for singly truncated samples when the number of unmeasured observations in the omitted portion is not known. The object of this paper is to study Cohen's problem when the number of unmeasured observations is known and to estimate the mean and variance of Type III populations from doubly and singly truncated samples with known truncation points when the number of unmeasured observations is (*i*) unknown for each tail, (*ii*) known separately for each tail, (*iii*) known jointly for the two tails.

I. THE PROBLEM

The Type III population may be written as

$$f(x)dx, \qquad \mu - \frac{2\sigma}{\alpha_3} \leqq x < \infty, \tag{1}$$

where

$$f(x) = \frac{c}{\sigma}\left[1 + \frac{\alpha_3}{2}\left(\frac{x-\mu}{\sigma}\right)\right]^{(4/\alpha_3^2)-1} e^{-(2/\alpha_3)(x-\mu)/\sigma}, \tag{2}$$

and

$$c = (4/\alpha_3^2)^{(4/\alpha_3^2)-1/2} e^{-4/\alpha_3^2} [\Gamma(4/\alpha_3^2)]^{-1}.$$

The mean, μ, and the variance, σ^2, of this population are unknown while in the first instance α_3, the third standard moment is supposed to be known. Let x_0' and $x_0'' = x_0' + R$ be the left and right truncation points respectively. Let there be n_1 unmeasured observations in the tail to the left of x_0', n_2 unmeasured observations in the tail to the right of x_0'' and n_0 measured observations in the sample range. We shall estimate the mean μ and the variance σ^2 in the three cases: (*i*) when n_1 and n_2 are not known, (*ii*) when n_1 and n_2 are known, (*iii*) when $n_1 + n_2 = N - n_0$ is known.

II. ESTIMATION BY THE METHOD OF MOMENTS

The truncated population may be written as

$$f(x')dx', \qquad 0 \leqq x' \leqq R, \tag{3}$$

where

$$f(x') = c\left[1 + \frac{\alpha_3}{2}(x' + \xi')\right]^{(4/\alpha_3^2)-1} e^{-(2/\alpha_3)(x'+\xi')}, \tag{4}$$

and

$$\frac{x - \mu}{\sigma} - \xi' = x', \qquad \frac{x_0' - \mu}{\sigma} = \xi', \qquad \frac{x_0'' - \mu}{\sigma} = \xi'', \tag{5}$$

$$\xi'' = \xi' + \frac{R}{\sigma}.$$

It is easy to see that $f(x')$ satisfies the differential equation

$$\frac{1}{f(x')} \frac{d}{dx'} f(x') = -\frac{\left(\frac{\alpha_3}{2} + \xi'\right) + x'}{\left(1 + \frac{\alpha_3}{2}\xi'\right) + \frac{\alpha_3}{2}x'}. \tag{6}$$

Case (*i*). To obtain the moments of the truncated population about x_0', we separate the variables in (6), multiply both sides of the resulting equation by x'^r and integrate over the range of the truncated population. Putting $r = 0$, 1, the first and second moments are given by

$$(7)\quad \begin{cases} \mu_1 = \sigma\left[-\xi' + \theta(Z_1 - Z_2) - \dfrac{\alpha_3}{2}\dfrac{R}{6}Z_2\right], \\ \mu_2 = \sigma^2\left\{1 + \xi'^2 + \theta\left(\dfrac{\alpha_3}{2} - \xi'\right)(Z_1 - Z_2) - \dfrac{R}{\sigma}Z_2\left[1 + \dfrac{\alpha_3}{2}\left(\dfrac{\alpha_3}{2} + \dfrac{R}{\sigma}\right)\right]\right\}, \end{cases}$$

where

$$(8)\quad \begin{cases} Z_1 = f(\xi')\left[\int_{\xi'}^{\xi''} f(t)dt\right]^{-1}, \\ Z_2 = f(\xi'')\left[\int_{\xi'}^{\xi''} f(t)dt\right]^{-1}, \end{cases}$$

and

$$(9)\quad f(t) = c\left[1 + \frac{\alpha_3}{2}t\right]^{(4/\alpha_3^2)-1} e^{-(2/\alpha_3)t},$$

$$(10)\quad \theta = 1 + \frac{\alpha_3}{2}\xi'.$$

Let

$$(11)\quad \nu_r = \sum_{i=1}^{n_0} (x_i - x_0')^r / n_0$$

be the rth moment of the sample of n_0 measured observations about x_0'. Equating the population and sample moments, the estimating equations obtained are

$$(12)\quad \begin{aligned} \nu_1 &= \sigma\left[-\xi' + \theta(Z_1 - Z_2) - \frac{\alpha_3}{2}\frac{R}{\sigma}Z_2\right], \\ \nu_2 &= \sigma^2\left\{1 + \xi'^2 + \theta\left(\frac{\alpha_3}{2} - \xi'\right)(Z_1 - Z_2) - \frac{R}{\sigma}Z_2\left[1 + \frac{\alpha_3}{2}\left(\frac{\alpha_3}{2} + \frac{R}{\sigma}\right)\right]\right\}. \end{aligned}$$

These equations in $\xi^{*\prime}$ and σ^* can be solved by Newton-Raphson method Here Z_1 and Z_2 can be calculated from Salvosa's tables [11]. Of course, μ^* is given by

$$(x_0' - \mu^*)/\sigma^* = \xi^{*\prime}.$$

Case (ii). The first two moments of the left tail about x_0' are

$$(13)\qquad \begin{aligned} m_1 &= -\sigma\left[\xi' + \theta\frac{n_0}{n_1}Y_1\right], \\ m_2 &= \sigma^2\left\{\theta - \left(\frac{\alpha_3}{2} - \xi'\right)\left(\xi' + \theta\frac{n_0}{n_1}Y_1\right)\right\}, \end{aligned}$$

while the first two moments of the right tail about x_0' are

$$(14)\qquad \begin{aligned} M_1 &= \sigma\left[-\xi' + \left(\theta + \frac{\alpha_3}{2}\frac{R}{\sigma}\right)\frac{n_0}{n_2}Y_2\right], \\ M_2 &= \sigma^2\left[1 + \xi'^2 + \left(\theta + \frac{\alpha_3}{2}\frac{R}{\sigma}\right)\left(\frac{\alpha_3}{2} - \xi' + \frac{R}{\sigma}\right)\frac{n_0}{n_2}Y_2\right], \end{aligned}$$

where

$$(15)\qquad \begin{aligned} Y_1 &= \frac{n_1}{n_0}f(\xi')\left[\int_{-2/\alpha_3}^{\xi'} f(t)dt\right]^{-1}, \\ Y_2 &= \frac{n_2}{n_0}f(\xi'')\left[\int_{\xi''}^{\infty} f(t)dt\right]^{-1}. \end{aligned}$$

Considering m_1, m_2 and M_1, M_2 as the contributions per unit observation of the left and right tails towards the first and second moments respectively, the modified moments of the total sample of $n_0+n_1+n_2$ observations about x_0' may be written as

$$(16)\qquad \begin{aligned} \nu_1' &= (n_1m_1 + n_0\nu_1 + n_2M_1)/(n_1 + n_0 + n_2), \\ \nu_2' &= (n_1m_2 + n_0\nu_2 + n_2M_2)/(n_1 + n_0 + n_2). \end{aligned}$$

The moments of the complete population about x_0' are given by

$$(17)\qquad \begin{cases} \mu_1' = -\xi'\sigma, \\ \mu_2' = (1 + \xi'^2)\sigma^2. \end{cases}$$

Equating the population moments to the corresponding sample moments, the estimating equations obtained are similar to (12) where Y_1 and Y_2 are substituted for Z_1 and Z_2 respectively.

Case (iii). The first two moments of the combined tail about x_0' are given by

$$g_1 = \sigma\left[-\xi' - \theta\frac{n_0}{N - n_0}(X_1 - X_2) + \frac{\alpha_3}{2}\frac{R}{\sigma}\frac{n_0}{N - n_0}X_2\right],$$

$$(18)\qquad g_2 = \sigma^2\left\{1 + \xi'^2 - \theta\left(\frac{\alpha_3}{2} - \xi'\right)\frac{n_0}{N - n_0}(X_1 - X_2) + \frac{R}{\sigma}\left[1 + \frac{\alpha_3}{2}\left(\frac{\alpha_3}{2} + \frac{R}{\sigma}\right)\right]\frac{n_0}{N - n_0}X_2\right\},$$

where

$$(19)\qquad X_1 = \frac{N - n_0}{n_0}f(\xi')\left[\int_{-2/\alpha_3}^{\xi'} f(t)dt + \int_{\xi''}^{\infty} f(t)dt\right]^{-1},$$
$$X_2 = \frac{N - n_0}{n_0}f(\xi'')\left[\int_{-2/\alpha_3}^{\xi'} f(t)dt + \int_{\xi''}^{\infty} f(t)dt\right]^{-1}.$$

The moments of the total sample of N observations about x_0' are given by

$$(20)\qquad \nu_1' = [n_0\nu_1 + (N - n_0)g_1]/N,$$
$$\nu_2' = [n_0\nu_2 + (N - n_0)g_2]/N.$$

Equating the sample moments to the population moments (17), we obtain estimating equations similar to (12), where X_1 and X_2 are substituted for Z_1 and Z_2 respectively.

III. MAXIMUM LIKELIHOOD ESTIMATION

Case (i). The likelihood function of the sample may be written as

$$(21)\qquad L_1 = \frac{K}{\sigma^{n_0}}\left[\int_{\xi'}^{\xi''} f(t)dt\right]^{-n_0} e^{-(2/\alpha_3)\sum_1^{n_0}\left(\frac{x''}{\sigma} + \xi'\right)} \cdot \prod_1^{n_0}\left[1 + \frac{\alpha_3}{2}\left(\frac{x''}{\sigma} + \xi'\right)\right]^{(4/\alpha_3^2)-1},$$

where

$$(22)\qquad x'' = x - x_0',\quad \xi' = \frac{x_0' - \mu}{\sigma},\quad \xi'' = \frac{x_0'' - \mu}{\sigma},\quad k = \text{constant},$$

and $f(t)$ is given by (9). Taking logarithms of (21) and differentiating partially with respect to ξ' and σ and setting to zero, the equations obtained are

$$(23)\qquad \begin{cases} \dfrac{1}{n_0}\left(\dfrac{4}{\alpha_3^2} - 1\right)\dfrac{\alpha_3}{2}S - \dfrac{2}{\alpha_3} + (Z_1 - Z_2) = 0, \\ -\dfrac{4}{\alpha_3^2} + \left(\dfrac{4}{\alpha_3^2} - 1\right)\theta\dfrac{S}{n_0} + \dfrac{2}{\alpha_3}\dfrac{\nu_1}{\sigma} + \dfrac{R}{\sigma}Z_2 = 0, \end{cases}$$

where

$$S = \sum\left[1 + \frac{\alpha_3}{2}\left(\frac{x''}{\sigma} + \xi'\right)\right]^{-1}, \tag{24}$$

and Z_1, Z_2, ν_1 given by (8) and (11) respectively. From these we obtain the maximum likelihood equations

$$\begin{cases} \sigma\left[-\xi' + \theta(Z_1 - Z_2) - \dfrac{\alpha_3}{2}\dfrac{R}{\sigma}Z_2\right] - \nu_1 = 0, \\ \dfrac{1}{n_0}\left(\dfrac{2}{\alpha_3} - \dfrac{\alpha_3}{2}\right)\sum_1^{n_0}\left[\theta + \dfrac{\alpha_3}{2}\dfrac{x'}{\sigma}\right]^{-1} - \dfrac{2}{\alpha_3} + (Z_1 - Z_2) = 0. \end{cases} \tag{25}$$

These equations in $\hat{\xi}'$ and $\hat{\sigma}$ can be solved by Newton-Ralphson method.

Case (ii). The likelihood function of the sample in this case may be put as

$$L_2 = \frac{K}{\sigma^{n_0}}\left[\int_{-2/\alpha_3}^{\xi'} f(t)dt\right]^{n_1}\left[\int_{\xi''}^{\infty} f(t)dt\right]^{n_3} e^{-(1/\alpha_3)} \cdot \sum_1^{n_0}\left(\frac{x''}{\sigma} + \xi'\right)\prod_1^{n_0}\left[1 + \frac{\alpha_3}{2}\left(\frac{x''}{\sigma} + \xi'\right)\right]^{(4/\alpha_3^2)-1}. \tag{26}$$

Proceeding as in case (*i*), the estimating equations obtained are similar to (25), where Y_1 and Y_2 are substituted for Z_1 and Z_2 respectively.

Case (iii). In this case the likelihood may be written as

$$L_3 = \frac{K}{\sigma^{n_0}}\left[\int_{-2/\alpha_3}^{\xi'} f(t)dt + \int_{\xi''}^{\infty} f(t)dt\right]^{N-n_0} e^{-(1/\alpha_3)}\sum_1^{n_0}\left(\frac{x''}{\sigma} + \xi'\right) \cdot \prod_1^{n_0}\left[1 + \frac{\alpha_3}{2}\left(\frac{x''}{\sigma} + \xi'\right)\right]^{(4/\alpha_3)-1}.$$

Proceeding as in case (*i*), the extimating equations obtained are similar to (25), where X_1 and X_2 are substituted for Z_1 and Z_2 respectively.

IV. THE SINGLY TRUNCATED CASE

The results for the singly truncated case can now be obtained in particular. In case truncation is on the left only, $x_0'' \to \infty$.

So that

$$
(27)\quad
\begin{cases}
\xi'' \text{ and } R \to \infty, \quad Z_2, RZ_2, X_2 \text{ and } RX_2 \to 0; \\
Z_1 \to \lambda(\xi'), \quad X_1 \to t(\xi'), \text{ where} \\
\lambda(\xi') = \left[1 + \dfrac{\alpha_3}{2}\xi'\right]^{(4/\alpha_3^2)-1} e^{-(2/\alpha_3)\xi'} \\
\qquad\qquad \cdot\left[\displaystyle\int_{-2/\alpha_3}^{\xi'} \left(1 + \frac{\alpha_3}{2}t\right)^{(4/\alpha_3^2)-1} e^{-(1/\alpha_3)t}dt\right]^{-1}, \\
t(\xi') = \dfrac{n_1}{n_0}\lambda(\xi').
\end{cases}
$$

Making relevant substitutions from (27) in case (*iii*) of Section II, we obtain the following equations of estimation by moments, when n_1, the number of unmeasured observations in the left tail, is known.

$$
(28)\quad
\begin{aligned}
\nu_1 &= \sigma[\theta t(\xi') - \xi'], \\
\nu_2 &= \sigma^2\left\{\theta + \left[\frac{\alpha_3}{2} - \xi'\right][\theta t(\xi') - \xi']\right\}.
\end{aligned}
$$

To solve these equations, we eliminate σ to obtain

$$
(29)\quad \frac{\nu_2}{\nu_1^2} = \frac{1}{\theta t(\xi') - \xi'}\left[\frac{\theta}{\theta t(\xi') - \xi'} - \frac{\alpha_3}{2} - \xi'\right].
$$

Corresponding to any ξ', the right-hand side of (29) can be evaluated from Solvosa's tables of areas and ordinates of the standardised Type III function. Hence, for a given sample $\xi^{*\prime}$ is determined from (29). And then

$$
(30)\quad
\begin{aligned}
\sigma^* &= \nu_1[\theta^* t(\xi^{*\prime}) - \xi^{*\prime}]^{-1}, \\
\mu^* &= x_0' - \xi^{*\prime}\sigma^*.
\end{aligned}
$$

Again, substituting in case (*iii*) of Section III, the maximum likelihood equations for this case are

$$
(31)\quad
\begin{aligned}
&\frac{1}{n_0}\frac{\alpha_3}{2}\left(\frac{4}{\alpha_3^2} - 1\right)S - \frac{2}{\alpha_3} + t(\xi') = 0, \\
&-\frac{4}{\alpha_3^2} + \frac{1}{n_0}\left(\frac{4}{\alpha_3^2} - 1\right)\theta S + \frac{2}{\alpha_3}\frac{\nu_1}{\sigma} = 0.
\end{aligned}
$$

It is easy to see that $\hat{\xi}'$ is given by

$$R(\hat{\xi}') = \frac{1}{n_0}\left(\frac{2}{\alpha_3} - \frac{\alpha_3}{2}\right)\sum\left\{\theta + \frac{\alpha_3}{2}\frac{x'}{\nu_1}\left[\theta t(\xi') - \xi'\right]\right\}^{-1}$$

(32)

$$-\frac{2}{\alpha_3} + t(\xi') = 0,$$

while

$$\hat{\sigma} = \nu_1[\hat{\theta}t(\hat{\xi}') - \hat{\xi}']^{-1}, \tag{33}$$

and

$$\hat{\mu} = x_0' - \hat{\xi}'\hat{\sigma}. \tag{34}$$

In case the number of unmeasured observations in the left tail is not known, we make relevant substitutions in case (*i*) of Section II and in case (*i*) of Section III and obtain equations by the method of moments and by the method of maximum likelihood respectively. The manner of solving these equations is the same as before.

V. ESTIMATION WHEN α_3 IS NOT KNOWN

We shall consider the singly truncated case when, n_1, the number of unmeasured observations in the left tail is known. The third moment of the left tail about x_0' is given by

$$m_3 = \sigma^3\left[\alpha_3 - 3\xi' - \xi'^3 - \theta\left(2 + \xi'^2 + \frac{\alpha_3^2}{2} - \frac{\alpha_3}{2}\xi'\right)\lambda(\xi')\right], \tag{35}$$

and the third modified moment, ν_3', of the total sample of n_0+n_1 observations about x_0' is given by

$$\nu_3' = (n_1 m_3 + n_0\nu_3)/(n_0 + n_1),$$

while the third moment, μ_3', of the population about x_0' is

$$\mu_3' = (\alpha_3 - 3\xi' - \xi'^3)\sigma^3. \tag{36}$$

Equating the first three moments of the population to the corresponding modified moments of the sample, the equations of estimation are (28) and

$$\nu_3 = \sigma^3\left\{\theta(\alpha_3 - \xi')\right.$$

(37)

$$\left. + \left[\theta t(\xi') - \xi'\right]\left[\frac{\alpha_3}{2}(\alpha_3 - \xi') + \xi'^2 + 2\right]\right\}.$$

To solve these equations, we estimate σ to obtain the following.

$$\frac{\nu_2}{\nu_1^2} = Q(\xi_1'\alpha_3) = \frac{1}{\theta t(\xi') - \xi'}\left[\frac{\theta}{\theta t(\xi') - \xi'} + \frac{\alpha_3}{2} - \xi'\right],$$

$$\frac{\nu_3}{\nu_1^3} = P(\xi_1'\alpha_3) = \frac{1}{[\theta t(\xi') - \xi']^2} \cdot \left\{\alpha_3\xi' + 2 + (\alpha_3 - \xi')\left[\frac{\theta}{\theta t(\xi') - \xi'} + \frac{\alpha_3}{2} - \xi'\right]\right\}. \tag{38}$$

From these equations we obtain $\xi^{*\prime}$ and α_3^* by Newton-Raphson method.

In cases, however, when truncation is on the right, the equations of estimation are

$$\begin{cases} \nu_1 = \sigma[\phi T(\xi') - \xi'], \\ \nu_2 = \sigma^2\left\{\phi + \left(\frac{\alpha_3}{2} - \xi'\right)[\phi T(\xi') - \xi']\right\}, \\ \nu_3 = \sigma^3\left\{\phi(\alpha_3 - \xi') + \left[\frac{\alpha_3}{2}(\alpha_3 - \xi') + \xi'^2 + 2\right][\phi T(\xi') - \xi']\right\}, \end{cases} \tag{39}$$

where

$$\phi = -\left(1 + \frac{\alpha_3}{2}\xi'\right), \tag{40}$$

and

$$T(\xi') = \frac{n_1}{n_0}\left[1 + \frac{\alpha_3}{2}\xi'\right]^{(4/\alpha_3^2)-1} e^{-(2/\alpha_3)\xi'} \cdot \left[\int_{t'}^{\infty}\left(1 + \frac{\alpha_3}{2}t\right)^{(4/\alpha_3^2)-1} e^{-(2/\alpha_3)t}dt\right]^{-1}. \tag{41}$$

For a given ξ', $T(\xi')$ can be calculated from Solvosa's Tables. The manner of solving these equations is the same as discussed above.

VI. EXAMPLES INVOLVING SINGLE TRUNCATION

(a) *Method of moments.* The method of estimating the parameters in this case can be best illustrated by considering the example given by Cohen (1950). This example was constructed by arbitrarily truncating

a sample distribution of the weights of 629 University of Washington Freshmen so that the record of observations less than 110.5 lbs. is missing. It is known that there are 15 observations in the omitted portion of the distribution. For the truncated distribution we have

$$n_0 = 614, \quad n_1 = 15, \quad x_0' = 110.5, \quad \nu_1 = 32.6596,$$
$$\nu_2 = 1359.2793, \quad \nu_3 = 67269.15.$$

Case (I). α_3 is known. In this case it is given that $\alpha_3 = 0.6$. The equations in ξ' and σ are

$$32.6596 = \sigma[\theta t(\xi') - \xi],$$
$$1359.2793 = \sigma^2\{\theta + (.3 - \xi')[\theta t(\xi') - \xi']\}.$$

The equation to determine ξ' is

$$1.27434 = Q(\xi') = \frac{1}{\theta t(\xi') - \xi'}\left[\frac{\theta}{\theta t(\xi') - \xi'} + .3 - \xi'\right].$$

For a first approximation for ξ', we read from Salvosa's Tables the value of ξ' such that the area of the standardised Type III curve from $-2/\alpha_3$ to ξ' is $15/629 = .023847$. Then a first approximation is $\xi' = -1.68$. Details of further computations for determining ξ' are given in the table below.

TABLE 1
COMPUTATIONS FOR $Q(\xi')$
$\alpha_3 = 0.6$

ξ'	$I_\theta(\xi')$	$f(\xi')$	$\lambda(\xi')$	$t(\xi')$	$\frac{1}{\theta t(\xi') - \xi'}$	$\frac{\alpha_3}{2} - \xi' + \theta(6)$	$Q(\xi')$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
−1.68	.023402	.089266	3.814460	.093187	.579300	2.267333	1.31347
−1.75	.017741	.072787	4.102756	.100230	.556295	2.314240	1.28740
−1.79	.015003	.064209	4.279744	.104554	.543949	2.341848	1.27385
−1.78	.015656	.066295	4.234479	.103448	.546984	2.334895	1.27715
−1.789	.015068	.064418	4.275153	.104442	.544251	2.341152	1.27417

Thus $\xi' = -1.789$. From (30), $\sigma^* = 17.775$ and $\mu^* = 142.30$.

Case (II). α_3 is not known. In this case, the equations in α_3 and ξ' are

$$1.27434 = Q(\xi, \alpha_3), \qquad 1.93101 = P(\xi', \alpha_3).$$

The third standard moment as obtained from the truncated sample may be used as a first approximation for α_3. Thus, $\alpha_3 = .7$ for a first approximation. A first approximation for ξ' is obtained on the assumption

that the area of the Type III function from $-2/\alpha_3$ to ξ' is 15/629. This gives $\xi' = -1.62$ from Salvosa's Tables. Detailed computations involved in the various steps to solve the equations by Newton-Raphson method are given in the tables below.

TABLE 2
COMPUTATIONS OF $Q(\xi', \alpha_3)$ AND $P(\xi', \alpha_3)$

α_3	ξ'	$I_\alpha(\xi')$	$f(\xi')$	$\lambda(\xi')$	$t(\xi')$	$\frac{1}{\theta(\xi') - \xi'}$	$Q(\xi', \alpha_3)$	$P(\xi', \alpha_3)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0.7								
	−1.60	.26153	.106606	4.076244	.099583	.608340	1.34910	
	−1.80	.010334	.054564	5.280046	.128992	.541205	1.27197	1.93773
	−1.79	.010891	.056727	5.208613	.127246	.544210	1.27523	
	−1.81	.009800	.052450	5.352041	.130750	.538236		1.92641
0.6								
	−1.78	.015656	.066295	4.234479	.103448	.546984	1.27715	1.94147
	−1.79	.015003	.064209	4.279744	.104554	.543949	1.27385	1.93003
	−1.789	.015068	.064418	4.275153	.104442	.544251	1.27417	1.93117

TABLE 3
DIFFERENCES OF ξ' AND α_3

α_3	ξ' from $Q(\xi', \alpha_3)$	ξ' from $P(\xi', \alpha_3)$	Difference
0.7	−1.793	−1.806	.013
0.6	−1.789	−1.789	.000

We thus find that $\alpha_3 = 0.6$ and $\xi' = -1.789$. Consequently, the estimates of μ and σ are the same as obtained in Case (I). The results obtained by Cohen (1950) for the case when the numbers of unmeasured observations is unknown, are shown in the table below.

TABLE 4
ESTIMATES OF THE PARAMETERS

Parameters	Values from complete sample	Estimates obtained by Cohen		Estimates obtained in this paper	
		α_3 known	α_3 unknown	α_3 known	α_3 unknown
α_3	0.59	0.60	0.71	0.60	0.60
ξ'	−1.805	−1.850	−1.878	−1.789	−1.789
σ	19.59	17.43	17.28	17.775	17.775
μ	142.25	142.75	142.95	142.30	142.30

(b) *Method of maximum likelihood.* The method of obtaining the maximum likelihood estimates is best illustrated by taking a random sample from the Type III population with $\alpha_3 = 0.6$, $\mu = 100.00$, $\sigma = 10.00$, truncated at 90.00. The random sample is

122.88	94.92	96.06	103.93	96.15
90.44	100.93	104.53	105.94	94.62
100.41	90.88	100.67	93.46	103.43
90.27	93.26	131.28	94.11	103.84
109.99	92.39	90.22	90.49	91.57
102.38	99.59	94.42	95.42	102.71

For the sample selected, we have

$$n_0 = 30, \quad n_1 = 5, \quad \nu_1 = 9.373, \quad \nu_2 = 172.74529, \quad x_0' = 90.00.$$

On the assumption that the area under the standardised Type III curve with $\alpha_3 = 0.6$ from $-2/\alpha_3$ to ξ' is $5/35 = .142857$, a first approximation for ξ' from Salvosa's Tables is given by $\xi' = 1.04$. Beginning with this value we finally obtain the following table.

TABLE 5
COMPUTATION OF $R(\xi')$

ξ'	S	$R(\xi')$
$-.75$	30.916744	$+.04939$
$-.80$	28.811959	$-.15082$
$-.76$	29.880896	$-.05288$

Hence, $\hat{\xi}' = -.755$ by interpolation.

From (33) and (34), $\hat{\sigma} = 9.8198$, $\hat{\mu} = 97.4139$. The estimates obtained by the method of moments are

$$\xi^{*\prime} = -.75, \qquad \sigma^* = 9.8773, \qquad \mu^* = 97.408.$$

The two methods appear to give estimates close to each other.

VII. AN EXAMPLE INVOLVING DOUBLE TRUNCATION

The manner of solving the estimating equations in this case will be illustrated by drawing a random sample of size 30 from the Type III population with parameters $\alpha_3 = 0.6$, $\mu = 100$, $\sigma = 10$, truncated at $x_0' = 85$ and $x_0'' = 115$. Only case (*i*) will be considered as the method is the same for other cases.

The sample drawn is

99.937	101.172	112.459	109.973	99.467
102.483	95.331	97.380	95.774	95.591
95.441	102.363	99.116	86.741	96.028
102.497	87.314	97.187	88.063	98.965
102.807	86.637	94.561	100.306	110.087
110.324	105.355	107.434	99.955	97.805

For the sample selected, we have

$$n_0 = 30, \qquad \nu_1 = 14.118433, \qquad \nu_2 = 246.292945, \qquad R = 30.$$

(a) *Method of moments.* The equations to be solved are

$$P(\xi', \sigma) = \sigma\left[-\xi' + (1 + .3\xi')(Z_1 - Z_2) - .3 = \frac{30}{\sigma} Z_2\right] - 14.118433 = 0,$$

$$Q(\xi', \sigma) = \sigma^2\left\{1 + \xi'^2 + (1 + .3\xi')(.3 - \xi')(Z_1 - Z_2) - \frac{30}{\sigma} Z_2\left[1 + .3\left(.3 + \frac{30}{\sigma}\right)\right]\right\} - 246.292945 = 0.$$

As a first approximation for σ, we use the sample standard deviation $s = 6.853$ and for ξ' the corresponding quantity calculated from the sample, viz., $-\nu_1/s = -2.06$. After several steps we obtain the table given below.

TABLE 6

DIFFERENCES OF σ AND ξ'

σ	ξ' from $P(\xi', \sigma)$	ξ' from $Q(\xi', \sigma)$	Difference
8.28729	−1.7960	−1.7963	+.0003
8.31025	−1.7917	−1.7913	−.0004

By interpolation we have, $\xi^* = -1.7942$, $\sigma^* = 8.2971$ and consequently $\mu^* = 99.8867$.

(b) *Method of maximum likelihood.* The equations to be solved are

$$(1°)\ \sigma\left[-\xi' + (1 + .3\xi')(Z_1 - Z_2) - .3 + \frac{30}{\sigma} Z_2\right] - 14.118433 = 0,$$

$$(2°)\ .101111 \sum\left[1 + .3\xi' + .3\frac{x'}{\sigma}\right]^{-1} - 3.33333 + (Z_1 - Z_2) = 0.$$

First approximations for σ and ξ' are the same as in (a).

After several steps we obtain the following table

TABLE 7
DIFFERENCES OF σ AND ξ'

σ	ξ' from (1°)	ξ' from (2°)	Difference
8.982036	−1.6739	−1.6752	+.0013
8.902077	−1.6871	−1.6860	−.0011

From this we have $\hat{\sigma}=8.9387$, $\hat{\xi}'=-1.6811$, $\hat{\mu}=100.0269$.

It may be noted that the two methods do not give the same estimates. These estimates, however, will be identical when $\alpha_3=0$, as shown by the author [6].

REFERENCES

[1] Cochran, W. G., "Use of IBM equipment in an investigation of the truncated normal problem," *Proceedings Research Forum, International Business Machines Corporation*, 1946, 40–3.

[2] Cohen, A. C., "Estimating parameters of Type III populations from truncated samples," *Journal of the American Statistical Association*, 45 (1950) 411–23.

[3] Cohen, A. C., "Estimating the mean and variance of normal populations from singly and doubly truncated samples," *Annals of Mathematical Statistics*, 21 (1950), 557–69.

[4] Cohen, A. C., "On estimating the mean and standard deviation of truncated normal distributions," *Journal of the American Statistical Association*, 44 (1949), 518–25.

[5] Des Raj, "On estimating the parameters of normal populations from singly truncated samples," *Ganita*, 3 (1952), 41–57.

[6] Des Raj, "On moments estimation of the parameters of normal populations from truncated samples," (to be published).

[7] Fisher, R. A., "Properties and applications of the functions," Introduction to *Mathematical Tables*, 1 (1931), British Association for the Advancement of Science.

[8] Hald, A., "Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point," *Skandinavisk Actuarietidskrift*, 32 (1949), 109–34.

[9] Lee, Alice, "Tables of Gaussian tail functions when the tail is larger than the body," *Biometrika*, 10 (1915), 208–15.

[10] Pearson, Karl, and Lee, Alice, "On the generalised probable error in multiple normal correlation," *Biometrika*, 6 (1908), 59–68.

[11] Salvosa, L. R., "Table of Pearson's Type III function," *Annals of Mathematical Statistics*, 1 (1930), appended.

[12] Stevens, W. G., "The truncated normal distribution," *Annals of Applied Biology*, 24 (1937), 815–52.

PROCEEDINGS

AMERICAN STATISTICAL ASSOCIATION
112TH ANNUAL MEETING

PALMER HOUSE, CHICAGO, ILLINOIS
DECEMBER 28, 1952

MINUTES OF THE ANNUAL BUSINESS MEETING

The meeting was called to order by Aryness Joy Wickens, outgoing President of the Association.

Report of the Committee on Elections

R. J. Eggert, Chairman of the Committee on Elections, reported to the membership the results of the balloting for officers for 1953. The following officers were elected:

President Elect	Herbert Marshall
Vice President	Rensis Likert
Directors (1953–55)	Wilfrid J. Dixon
	Margaret J. Hagood
Representative at Large (1953–54)	Kenneth Miller
District Representatives	
Northeastern District	David Votaw, Jr.
Eastern District	A. J. Jaffe
Southeastern District	Ezra Glaser
North Central District	William Madow
South Central District	John Stockton
Western District	Harry Schwartz

Nominations for District Representatives

Mrs. Wickens presented to the membership some of the difficulties attached to the nomination of district representatives. She stated that the matter would be presented to the incoming Board and Council. However, she remarked that it was up to the membership to make these nominations and that heretofore the nominations were late in arriving at the Secretary's office. Since this was a matter of interest to the whole membership, Mrs. Wickens urged that these nominations be forwarded to the Secretary's office early in the year.

The Report of the Board of Directors for 1952

Morris Hansen read the report to the membership. It was moved that the report be received and approved. The motion was carried.

Secretary-Treasurer's Report for 1952

Samuel Weiss read the report of the Secretary-Treasurer to the membership. It was moved that the report be accepted and approved. The motion was carried.

Coming Annual Meetings

Mrs. Wickens asked Mr. Weiss to present to the membership the dates and places that had been decided upon for future annual meetings of the Association. They are as follows:

Year	*Place*	*Hotel*	*Date*
1953	Washington, D. C.	Shoreham	December 27–30
1954	Montreal	Mount Royal	September 10–13
1954	San Francisco	(This is a regional meeting—hotel and dates not yet chosen.)	
1955	New York City	Biltmore	Late in December

Resolutions

Mrs. Wickens called upon Ralph Watkins to present the following resolutions to the membership:

1. Resolution regarding the Program Committee.
   RESOLVED that the officers and members of the American Statistical Association express deep appreciation for the excellent program prepared by the members of the Program Committee under the leadership of Alfred Watson, Chairman, and Edward Bloom, Secretary.

The resolution was approved.

2. Resolution regarding local Arrangements Committee and Chicago Chapter.
   RESOLVED that the officers and members of the American Statistical Association express their profound appreciation to the Local Arrangements Committee under the Chairmanship of Wesley Mitchell and to all of the individual members of the Chicago Chapter for their outstanding work and hospitality in connection with the arrangements for the 112th annual meeting of the Association.

The resolution was approved.

3. Resolution regarding the retirement of Sylvia Weyl.
   WHEREAS, Sylvia C. Weyl has ably and faithfully served the Association in the capacity of Executive Assistant and Editor of *The American Statistician*, the former from 1944 to July 1952, and the latter from 1947 to December 1952, and has now retired from these positions;
   THEREFORE, BE IT RESOLVED, that the members of the American Statistical Association in annual meeting assembled do hereby express great appreciation of the outstanding service rendered by Mrs. Weyl during her terms of office, and hereby express our recognition of her valuable assistance so competently performed over this long term of years.

The resolution was approved.

Mrs. Wickens turned the chairmanship of the meeting over to Professor Cochran, the president of the Association for 1953.

New Business

Professor Cochran asked if there was any new business to be discussed. It was moved that the membership consider a time other than Christmas week for the 1956 annual meeting of the Association. The motion was carried. The meeting was adjourned.

Report of the Board of Directors for 1952

During the calendar year just ended, the Association continued its policy of strict economy, added moderately to its surplus, registered a significant increase in membership and expanded the scope of its manifold activities.

Program Committee

Organized around the theme, Applications and Uses of Statistics, the Annual Meeting was planned and directed by a Program Committee chaired by Alfred Watson with Edward Bloom as secretary. The section on Business and Economic Statistics under the chairmanship of Donald Riley, the section on the Training of Statisticians under Philip McCarthy, the Biometrics section under Alexander Mood, as well as the Committee on Statistics in the Physical Sciences under the chairmanship of John Tukey and the Committee on Statistics in the Social Sciences under Conrad Taeuber all contributed substantial sections of the Annual Meeting Program.

Bureau of Mines Survey

The Board is happy to report that the Survey of the Statistical Program of the Bureau of Mines under the able direction of J. E. Morton, advised by a committee consisting of Raymond Bowman, Ralph Watkins and Clarence Long, was completed last summer. The Report has been submitted to the Bureau of Mines for its consideration and action.

Other Committees

The Association's advisory committee to the Bureau of Labor Statistics was engaged during 1952 in the task of cooperating in the preparation of the new consumers' price index. The committee to advise the Bureau of the Budget on statistical policy reports started its work during the calendar year.

Business and Economic Statistics Section

This Section drafted its charter last year, which was thereupon submitted to and accepted by the Board which modified it slightly and referred it back to the Section for its consideration. This Section took the leadership in holding two conferences during 1952, one in Illinois and one in Pennsylvania. The Pennsylvania meeting was organized in conjunction with The University of Pennsylvania, Wharton School of Business Administration on the topic, "The Role of Statistics in Business Planning and Control." The Middle Western meeting was organized with the University of Illinois around the topic, "Minimizing Business Risks." All institutional members of the Association and all members of the Section were invited.

Tulsa Chapter

The Association welcomes its newest chapter, organized in Tulsa, Oklahoma, in 1952. The proposed Constitution of the group has already been approved and its Charter granted.

The American Statistician

After having served over five years as Editor of *The American Statistician*, the Association's news publication, Sylvia Weyl asked to be relieved from her assignment. Mrs. Weyl was the first editor of *The American Statistician* and from its inception in August of 1947 she made *The American Statistician* a medium for the discussion of non-theoretical matters of concern to the statistical profession, including the aims and status of the profession, discussions of statistical

institutions and non-theoretical expository articles on current statistical activities and projects. Under Mrs. Weyl's editorship, a substantial number of articles dealing with the training of statisticians, the needs and activities of business statisticians and descriptions and discussions of practical aids to working statisticians have appeared. The publication has filled the void between the technical professional discussions of statistical theory which appear in our *Journal* and the simple listing of program notes and announcements which used to be carried in the Association's *Bulletin*. Mrs. Weyl did a remarkable job of producing a readable and useful publication of a sort not previously undertaken by professional societies.

An *ad hoc* subcommittee of the Committee on Publications was designated to search for a new editor. Pursuant to its recommendations, *The American Statistician* will be edited under the sponsorship of the Wharton School of Business of the University of Pennsylvania. Dr. Raymond Bowman has agreed to give the necessary guidance and supervision; Almarin Phillips of his staff will serve as Editor.

JASA

At the request of W. Allen Wallis, Editor, and at the recommendation of the Secretary, the Board approved an expanded budget for the *Journal of the American Statistical Association* during the year just ended. Size and content of the periodical were expanded, pages printed in 1952 were 20% greater than in 1951.

Constitution

The Committee on the Constitution submitted a revised version to the Board of Directors. The proposed revisions include a number of changes of a noncontroversial character designed to permit greater ease and flexibility of operations. After careful consideration, the Board approved the Committee's proposals. The Constitution in its revised form is being submitted to the membership for approval.

Liaison

The Board is desirous of increasing membership participation in the direction of the Association's affairs. With this in view, it was decided to send minutes of all its meetings to the chairmen of sections and of standing committees, as well as to the members of the Board and Council. In order to ensure clearer presentation of the various opinions and viewpoints expressed at Board meetings, the Board at its last meeting decided to appoint a special rapporteur to prepare a summary of its views which will supplement the formal minutes.

Report of the Secretary-Treasurer for 1952

During 1952, the Secretary's office continued to devote its efforts to extremely careful management. Through rigorous economy the Association completed its third successive year in the black. During 1952 the increase in the Association's surplus resulted mainly from strict financial control rather than from membership gains.

By the end of 1951, it had become apparent that Association membership had remained virtually static for a period of three to four years. Accordingly, the

Secretary's Office undertook a special membership drive last year, circularizing by letter a large number of mathematicians, economists, agricultural statisticians and others whose professional activities brought them into close relationship with the work of the Association. This campaign was interrupted last summer, but will be resumed in 1953. It is encouraging to note that the Association has, for the first time since 1948, shown a significant growth in membership.

The number of members at the end of 1951 was 4,356. During 1952 about 350 members will be dropped because of resignation, death, or non-payment of dues. New members in 1952 totalled 666. Thus the Association starts the year 1953 with 4,655 members. The net membership growth in 1952 was 300 members—the first net increase in 5 years.

Expenses for 1952 were budgeted at $48,666. The actual expense for the year was $47,529.49—a drop of over $1,000. Although the increase in the cost of publications was about $300 over the 1952 budget, certain other expenses were kept under budget level by saving wherever possible.

While the saving in any particular item was not large in itself, the combined total was just about sufficient to offset the rise in the cost of publications. The actual income for 1952 was $54,061. This is about $700 under the budget figure. The loss is primarily due to the decrease in the sales of back issues of the *Journal*. The actual net income for 1952 is $6,531 which provides for a substantial increase to the Association's surplus. Thus the Association will start the year 1953 with $16,965 in surplus funds.

Office Organization

As a result of the resignation of Sylvia Weyl, who had been Executive Assistant to the Secretary since the beginning of 1944, it was necessary to reorganize the Secretary's office. With the agreement of the liaison committee of the Board of Directors, Edgar Bisgyer was appointed to serve in a managerial capacity in the Secretary's office. Mrs. Weyl, who in addition to her administrative work had been editor of *The American Statistician* since it was founded in 1947, agreed to continue to edit that publication until the Board of Directors and the Secretary were able to find a suitable person to assume the editorship. The transfer of editorial function from the Washington office to the University of Pennsylvania campus will thus be carried out in an orderly fashion and without loss of efficiency.

Financial Recommendations

The Treasurer's Report, which is printed separately, shows the strengthened financial position of the Association and emphasizes that 1952 was the third successive year of accruing surplus. In this connection, the Secretary has recommended to the Board of Directors that the Association plan to build up a surplus equivalent to one full year's income before entering into any major expansion of its activities. It is proposed that approximately $3,000 be added to surplus each year until this objective is attained. With this in mind the budget for 1953 has been calculated to provide a net income of slightly over $3,000. The proposed income is budgeted at $50,650, while expense has been calculated at $47,612. While expense is calculated as closely as possible, income is figured very conservatively and actual income figure for 1953 may be expected to be somewhat larger than that listed in the budget.

To the Board of Directors of
American Statistical Association.

I have examined the attached financial statements of American Statistical Association relating to the year ended December 31, 1952. My examination was made in accordance with generally accepted auditing standards and, accordingly, included such tests of the accounting records and such other auditing procedures as were considered necessary in the circumstances.

The recorded cash receipts for the year were traced to the deposits shown on the bank statements and the amounts for dues and subscriptions were tested with the membership and subscription records. The paid checks were inspected and related vouchers tested in support of cash disbursements for the year. The bank balances were reconciled with amounts reported directly to me by the depositories and the cash on hand at December 31, 1952, was verified by inspection. I did not check the membership and subscription records in detail or make any independent verification of the inventory of old *Journals*, the office records of which are based, in part, on data assembled in prior years.

In accordance with a resolution of the Board of Directors, the expense incurred in publishing a directory, distributed to the membership in 1951, is being spread over a three-year period although such costs would appear to be applicable primarily to the year 1951. The amount deferred at December 31, 1952, aggregated $831.86.

In my opinion, the accompanying statements present fairly the position of American Statistical Association at December 31, 1952, and the results of its operations for the year, in conformity with generally accepted accounting principles applied on a basis consistent, except as mentioned in the preceding paragraph, with that of the preceding year.

JAMES G. JESTER

American Statistical Association
Balance Sheet

Assets

	December 31, 1952	December 31, 1951
Cash in banks and on hand	$39,503.14	$28,618.28
Accounts receivable	1,099.58	1,385.48
Investment in United States Savings Bonds, Series G, due 1962, at cost	3,100.00	3,100.00
Inventory of old *Journals*, at approximate cost	1,909.85	1,625.85
Inventory of monograph on Acceptance Sampling, at cost	233.93	452.54
Inventory of emblems, at cost	463.50	588.00
Furniture and fixtures, at cost less depreciation	2,192.42	2,026.99
Deferred charges:		
Deferred Membership Directory expense	831.86	2,131.86
Other	1,007.68	723.48
	$50,342.15	$40,652.48

Liabilities

Accounts payable	$ 5,930.38	$ 4,811.46
Deferred income (collections applicable to subsequent year):		
Dues	$16,179.00	$13,618.00
Subscriptions	5,524.48	4,379.60
Other	2,865.16	4,365.67
	$24,568.64	$22,363.27
Net Worth:		
Life membership reserve	$ 2,877.79	$ 3,044.10
Surplus, per statement	16,965.34	10,433.65
	$19,843.13	$13,477.75
	$50,342.15	$40,652.48

American Statistical Association
Statement of Income and Surplus Accounts

	Year ended December 31, 1952	1951
Income:		
Dues—current year	$37,101.00	$33,358.00
Dues—prior year	184.00	400.00
Life membership income	166.31	84.40
Subscriptions	9,971.35	9,433.15
Advertising	1,594.47	2,251.13
Journal sales, less cost of sales	1,079.14	1,907.03
American Statistician sales	82.65	158.65
Acceptance Sampling monograph, less cost of sales	24.39	210.60
Mailing list income	704.17	683.34
Biometrics sales	453.25	490.80
Interest income	541.37	395.40
Reimbursement of overhead expenses: Bureau of Mines project	2,000.00	1,500.00
Miscellaneous	159.08	702.63
Total income	$54,061.18	$51,575.13
Expense:		
Journal—printing, mailing and reprints	$12,858.57	$ 9,997.19
Salaries and wages	14,716.81	14,817.96
American Statistician	5,810.75	4,815.68
Membership Directory—apportioned expense	1,300.00	1,233.50
Annual meeting expense	550.72	141.29
Rent	2,400.00	2,400.00
Office supplies, printing and mimeographing	2,434.54	1,659.39
Postage	1,163.45	1,911.95
Promotion expense	702.30	2,375.00
Telephone and telegraph	563.71	482.11
Travel and committee expense	2,001.67	1,196.52
Depreciation of furniture and equipment	555.87	510.86
Accounting services	970.00	970.00
Miscellaneous	1,501.10	1,947.46
Total expense	$47,529.49	$44,458.91
Excess of income over expense for the year	$ 6,531.69	$ 7,116.22
Add: Surplus account at beginning of year	10,433.65	3,317.43
Surplus account at end of year	$16,965.34	$10,433.65

BOOK REVIEWS

Wesley Clair Mitchell: The Economic Scientist. *Arthur F. Burns, editor.* New York: National Bureau of Economic Research, Inc., 1952. Pp. viii, 327. $4.00.

See the article by Adolf A. Berle, pp. 169–175 in this issue.

An Introduction to Scientific Research. *E. Bright Wilson* (Theodore William Richards Professor of Chemistry, Harvard University). New York: McGraw-Hill Book Company, 1952. Pp. v, 365. $6.50.

E. L. LEHMANN, *University of California (Berkeley)*

THIS volume has as its primary concern the experimental and observational aspects of science, and its main purpose is to provide a collection of general principles and specific methods applicable to the planning and analyzing of scientific work of this kind. Under these circumstances it seems quite natural (at least to a statistician) that the concept of randomness pervades the book, and that of the eight principal chapters, five are of an essentially statistical character. Taken together, these form an introduction to the concepts and methods of modern statistics.

Any book covering such wide territory faces a twofold danger. It may get lost in generalities which are of no real use to the practicing scientist. Alternatively, it may become a collection of rules and recipes, which are not only tedious but also likely to be misused if they do not grow out of an understanding of the basic concepts and are not set within a framework of theory that indicates the necessary limitations. The author manages to avoid this double pitfall by putting considerable emphasis on the conceptual side of the subject, and by developing the techniques as examples of the general approach rather than in the dogmatic way one so frequently encounters.

In setting forth the most important statistical notions, such as those of hypothesis testing, confidence intervals, factorial experimentation, and randomization, the author uses the device of first introducing a concept only qualitatively in a natural experimental setting. An analytical development is usually deferred until a later chapter and frequently the treatment of more specific problems at a still later point provides an opportunity for reiteration of the fundamental principles and further clarification of the conceptual difficulties. This gradual approach saves a reader unfamiliar with the subject from being overwhelmed by so many new ideas.

The development of the main statistical theories is carried out in Chapters 4 (design of experiments, 33 pages), 8 (analysis of experimental data, 63 pages) and 9, (errors of measurement, 45 pages), where numerous specific techniques are also given, such as testing and estimation in binomial, normal and Poisson distributions, testing for goodness of fit, analysis of variance, control charts. Many further methods are mentioned only briefly but always concretely, as for example run tests, sequential probability ratio tests, and tolerance limits.

Closely related to the above material is Chapter 7 (classification, sampling, and measurement, 16 pages) in which the concept of a sample is clarified, and scientific induction is discussed as the method of inference from a sample to the population from which it was drawn. A formal discussion of probability is postponed until Chapter 10 (probability, randomness, and logic, 22 pages), where it is defined in terms of a random sampling process and where the addition and multiplication rules are derived. The author here also discusses the problem of how to utilize prior information, and in this connection gives a brief account of the subjective point of view as presented by Jeffreys.

As the author states, the book is intended primarily for "students beginning research and for those more experienced research workers who wish an introduction to various topics which were not included in their training." While a number of derivations are sketched that require differential arguments, such as those of the t- and χ^2-distributions, these are usually given at the end of a chapter, so as not to interrupt the main development. Consequently, the book has much to offer even to a reader not familiar with the calculus. Of great help are the many examples, some particularly interesting ones being taken from the physical sciences, and the detailed discussion of the many pitfalls that threaten the unwary.

If fault must be found, the reviewer would have preferred a clearly stated general definition of mathematical expectation and variance to the one given here by implication. He also missed a statement of the additivity of expectations and other results of this type, which one feels might have been included in chapter ten. These criticisms, however, view the book as a text in statistics, which it is not meant to be. Also, they are concerned solely with the mathematical aspects of the theory. With regard to the meaning, intent and possibilities of statistical methods, the reviewer knows of no clearer or, on its level, more useful account.

Finally, it should be mentioned that in addition to the five chapters reviewed here the book contains five chapters of a general character (Chapter 1, choice and statement of a research problem, 9 pages; Chapter 2, searching the literature, 11 pages; Chapter 3, elementary scientific method, 15 pages; Chapter 6, execution of experiments, 24 pages; Chapter 13, reporting the results, 10 pages); and three important technical chapters (Chapter 5, design of apparatus, 58 pages; Chapter 11, mathematical work, 29 pages; Chapter 12, numerical computations, 22 pages).

Statistical Theory in Research. *R. L. Anderson* and *T. A. Bancroft.* New York: McGraw-Hill Book Company, 1952. Pp. xix, 399. $7.00.

CHURCHILL EISENHART, *National Bureau of Standards*

THIS volume treats in detail an amazingly large number of topics in statistical theory and methodology, with careful attention to the assumptions and mathematics upon which they are based, and illustrates their applica-

tion through numerous worked examples in the text, plus a large number of well chosen exercises at the end of each chapter. Each chapter has its own list of "References cited," to which explicit reference is made at appropriate points in the text. These are well chosen, as are the supplemental items listed as "other reading." Together they will enable the reader (or teacher) to gain (give) fuller information on topics treated incompletely, or only mentioned, in the text.

Written to fill a need expressed by "many research workers" for "a convenient reference book on statistical theory pointed to research problems, which could be used in conjunction with their books on general statistical methods, experimental design, and survey sampling," this book should be well received by students and research workers in the fields of agricultural, biological, and sociological research who (1) have a "good background in differential and integral calculus" and (2) have had some first-hand experience with conducting and interpreting experiments in one or more of these fields. For these, it will serve to explain the mathematical bases of the various statistical techniques with which they have become acquainted through their experimental work, or from elementary "practical" courses in statistical method, or both. Students and research workers in the physical sciences and engineering, on the other hand, will find the material covered herein somewhat less familiar, and its scope somewhat less satisfactory—some statistical principles and techniques of importance in these fields are only touched upon briefly here, and others are not even mentioned.

Viewed as a textbook, this volume is really two books in one. Part I, entitled "Basic Statistical Theory," and the first four chapters (Chaps. 13–16) on regression analysis in Part II, are suitable for a one-year course in mathematical statistics taught, for example, in a department of mathematics or statistics, and can be understood by a student with the mathematical maturity of a mathematics major even though he may have had no first-hand experience with experimentation. These same chapters will be more difficult for the experimental scientist; here his practical intuition born of experience will serve him less well.

Part II is entitled "Analysis of Experimental Models by Least Squares." Experimental scientists in agriculture, biology, and social science who have had some experience with analysis of variance and modern experimental arrangements will be much more at home here in spite of the fact that the complexity of the algebraic manipulations involved will give them plenty of opportunity to exercise their mathematical skills. Here their intuition derived from experience will serve them well, and I believe that they will fully appreciate the import of the mathematical details even though they may be unable to reproduce them without aid from the book itself. On the other hand, the student who approaches the subject from the mathematical viewpoint cannot in my opinion fully appreciate the material presented in Part II until he has gained some practical experimental experience, no matter how adept he becomes at the mathematics.

The physical scientist and engineer without previous training or experience in analysis of variance and the Fisherian principles and techniques of experiment design will be in much the same boat—they will have the mathematics but not the viewpoint needed to fully appreciate these chapters. Before tackling Part II, they will do well to gain the requisite perspective by reading, for example, Chapters 4 (design of experiments) and 8 (analysis of experimental data) in *An Introduction to Scientific Research*, by E. Bright Wilson, Jr. (see preceding review).

In the preface, the authors "welcome suggestions on methods of improving this book both from a reference and from a text standpoint, without adding materially to the complexity of the material or the length of the book." My own personal feeling is that a revision of this volume should not be attempted at all. Instead, Part II, of which Bancroft contributed the first four chapters and Anderson the remaining nine, should be split off from Part I, and issued separately as a monograph on "Analysis of Experiments by Least Squares, and the Method of Variance Components." Part II is a very compact and lucid treatment of the mathematics underlying these distinctly different aspects of "analysis of variance and covariance;" is filled with up-to-the-minute information on these topics; and contains in the last chapter a useful "summary of needed research." This material forms a very useful supplement to technique books on analysis of variance and the design of experiments, and in the form of a separate volume could and should be kept up-to-date as such.

Except for isolated sections, Part I is essentially a compact treatment of statistical theory appropriate to measurement data (in contrast to enumeration data, and rank-order data). It might well be expanded a bit to make it a fairly comprehensive monograph on this phase of statistics. As it stands, it is a very readable and yet compact coverage of the elementary theory of probability, univariate and multivariate distributions, mathematical expectations and moments (including moment-generating functions and cumulants), transformation of variables and derivation of sampling distributions of statistics, and orthogonal linear functions; together with point estimation (from the Fisherian viewpoint), interval estimation and tests of statistical hypotheses (from the Neyman and Pearson viewpoint). To round it out it needs a discussion of the concept and techniques of "statistical control" in relation to measurement processes, a compact discussion of the various different "straight-line situations" of importance to the measurement analyst, something on "the law of propagation of error", and a bit on sequential analysis.

It is my opinion that what is needed today is not more books that attempt to cover a wide (and for the most part the same) selections of material in detail, but rather, some books that provide a panoramic view of a large fraction of existing statistical theory and methodology, and others that are comprehensive monographs on some single phase in which the author is particularly well grounded.

Statistics for Sociologists. Revised Edition. *Margaret Jarman Hagood* and *Daniel O. Price.* New York: Henry Holt and Company, 1952. Pp. xii, 575. $5.75.

A. W. MARSHALL, *The Rand Corporation*

THIS is a revised version of a book published in 1941 by the senior author. The revisions and additions include (1) new chapters on sampling in social surveys and on applications of factor or component analysis in sociological research, (2) the introduction of the work of Guttman into the chapter on indexes and scales, (3) new, up to date reading references, (4) substitution of current illustrative materials in almost all examples, (5) reduction of space devoted to computational procedures, (6) systematic collection of the most useful formulae hitherto scattered throughout the text at the end of the text and, (7) deletion of Part V of the earlier work, which contained material on birth and death rates and the construction of life tables. All of these changes have undoubtedly added to the general usefulness of the book.

This book represents an attempt to write an essentially non-mathematical text, for a group of students who would not understand a mathematical one, in order to convey to them "all of the basic statistical methods which have been used in sociological research, and some of the newer ones which have not as yet found wide application in sociological fields." This same aspiration guided the preparation of the first edition 11 or 12 years ago. As then, books of this type for this particular audience raise very serious pedagogical problems. This book probably succeeds as well as any book could in making some progress toward the above goals and so long as there exists a demand for texts which make such an attempt it should prove useful.

The reviewer's private fantasy life on how to teach statistics to novices, in some substantive field such as sociology, contains two elements:

(1) A good supply of first rate papers by the leaders in the substantive field of study which contain well thought out applications of statistics to important problems.

(2) A textbook of the general type being reviewed here; i.e., one that is essentially an introduction to statistical tools of the trade, but with much more emphasis on first principles and the appropriate cautions to be observed in the use of the tools. In particular this model text would stress the logic of model construction and analysis of the structure of the population from which samples are assumed to be drawn. In addition it would have directions for the discovery of cases where sociologists or others ought to yell for help from mathematical statisticians, with perhaps some indications as to what kinds of help are probably forthcoming. (The importance of this latter problem is very much overlooked. In the course of the reviewer's collaboration with sociologists, when the problem they wished to deal with was completely stated, the statistical problems involved were very frequently not textbook cases.)

This model text would be slanted toward the needs of those students with serious research ambitions and to that extent is not directly comparable with

the present work and no unfair comparisons should be indulged in. The demand and audience for such a model text (if indeed the above model text is truly optimal) has not really developed in sociology, nor is it possible to satisfy condition (1) above. The authors themselves "regret that the progress in the application of statistics to sociological research since 1941 has not provided a basis for a more complete substitution of new titles" for older references for supplementary reading.

Given their audience, the authors have written a book which is, if anything, a little above the average abilities of those who will read it. As the paragraphs immediately above suggest this is bias in the right direction. The book is very sound on those problems it does treat and with close reading no one should be given any wrong ideas about the application of statistics to sociological problems.

Statistical Quality Control. Second Edition. *Eugene L. Grant.* New York: McGraw-Hill Book Company, 1952. Pp. xvi, 557. $6.50.

C. C. Craig, *University of Michigan*

Since two very good reviews of the first edition of this widely used textbook appeared in this *Journal*, Vol. 42 (1947), pp. 180–184, it is appropriate first to note the more important changes in the new edition.

The most extensive additions are in the area of acceptance sampling. Since the appearance of the first edition, the Military Standard 105A acceptance sampling tables for inspection by attributes have superseded the Army Service Forces tables. Professor Grant has drawn on the fine study he and Lorber made of the new tables to replace the section on the older tables with a very good and thorough discussion of the new ones. The clear explanations of the use of the tables, of their relation to their predecessors, and of their more important characteristics should be welcomed by many present and potential users. Another new feature of importance is the introduction of acceptance sampling by variables. The use of known-sigma plans for one- and two-sided specifications and of unknown-sigma plans for one-sided specifications is explained and relevant tables are provided. The $\overline{U}$ and Q tests of Schwartz and Kaufman are also included as is Shainin's Lot Plot plan, the latter in a rather brief and non-committal fashion. The reviewer was disappointed in finding no critical discussion of Shainin's widely used plan. Though group sequential plans come in for repeated mention and are explained, the reviewer was also disappointed with the very brief treatment given item-by-item sequential plans. An example is given of the calculation of an attributes plan with the very minimum of discussion. Surely the importance of sequential analysis for applications merits more attention.

There is a good deal of revision in other sections of the book, too. The section on the cost aspects of quality decisions has been somewhat expanded. Worthy of particular mention are: (1) the use of the exact value instead of the large sample approximation to the standard deviation of the sample stand-

ard deviation for small samples from a normal universe; (2) the addition of a brief discussion of the use of the normal frequency function as an approximation to the binomial; (3) a modification in the handling of varying sample units on *c* charts; and (4) a discussion of the application of control charts to clerical work. These are all improvements. But of greater importance is the fact that the new edition contains 302 problems as compared with 145 in the first edition. Perhaps the most difficult task in writing a book of this kind is to provide enough good problems; Grant deserves special commendation on this point.

There are also points regarding which the reviewer feels mildly critical. First, he does not agree that one loses any significant figures when he subtracts the square of the mean from the mean of squares in calculating a variance. The text appears to give the impression that one sacrifices something in return for the convenience of the "short-cut" method. Second, the reviewer would have liked to see considerable emphasis on the fact that setting limits on an $\overline{X}$ chart by use of an $\overline{R}$ calculated from R's whose chart shows lack of control is at best only tentative. Third, because it is so hard to make the point really stick, he would have been pleased by some real hammering at the fact that deleting points out of limit lines to set revised limits is only paper work unless the related assignable causes are actually removed. Fourth, on page 98 the idea of a confidence interval is introduced, though not so designated, at first correctly, and then marred by "a less precise but more common interpretation" which incorporates the common first misconception of this notion. In fact the reviewer believes that no book on statistical methods, even at the most elementary level, should fail to give considerable attention to this important and clarifying concept. Finally, on page 412 it seems that some attempt should have been made to explain the difference in origin between the sample standard deviation, as used elsewhere in the book, and the square root of the unbiased estimate of the variance and it probably should be stated that the latter is also a biased estimate of the standard deviation of a normal universe.

It should be remembered that this is a pre-calculus text and written for the student either in school or in a shop who has the very practical aim of making statistical quality control methods work. The reviewer has had a good deal of experience with the first edition and he is convinced that the second edition only keeps this book considerably the best in its field. As a text for engineering students who have had a year of the calculus one could wish that there were more relevant mathematical material and that, say, the first three chapters on $\overline{X}$ and R charts had been more concisely written. But the engineering point of view in this book is authentic and Grant keeps his feet firmly on the ground. This, in combination with its clear and systematic exposition, its wealth of examples, and its numerous problems makes its merits far outweigh the few imperfections noted.

Statistical Theory with Engineering Applications. *A. Hald.* New York: John Wiley and Sons, Inc., 1952. Pp. xii, 783. $9.00.

G. J. LIEBERMAN, *Stanford University*

ACCORDING to the author, the aim of this book is "first and foremost to furnish the reader with simple and practical methods—which can be understood and applied by non-statisticians—for the handling of the majority of the problems which occur in everyday work." Hald has not only accomplished this, but, in addition, has provided the statistician with a ready reference book containing an abundance of interesting examples drawn from the engineering fields. Furthermore, this book will be useful as a supplementary text in many of the introductory courses in mathematical statistics, inasmuch as the treatment of the topics assumes a minimum knowledge in mathematics of the calculus.

The book devotes chapters to the basic statistical topics including the calculus of probabilities, distribution theory, limit theorems, analysis of variance, design of experiments, regression, and correlation. Hald's treatment of most of these topics is excellent. For example, for an elementary text, his axiomatic approach to probability theory is quite refreshing. His handling of degrees of freedom as related to linear restraints on the random variables is much more mature than that found in the usual texts.

For the practical man, the oustanding achievement of the book is the very complete treatment of the usual standard techniques in the handling of data. For example, the chapter on Regression contains a section on solving the Normal Equations by the Doolittle method. Part of a chapter is devoted to criteria for rejection of outlying observations. Almost any topic from a treatment of the Central Limit Theorem to a section on Punched Cards can be found in the book, together with illustrative examples. From a practical point of view, *Statistical Theory* is far superior to anything that has been written to date.

The only criticism of the book is that it is somewhat incomplete from a theoretical point of view. Although Hald states that "a general exposition of the theory of statistics lies outside the scope of the present publication," the reviewer feels that because the author presents a great deal of theoretical as well as applied statistics, certain topics should have been included. There is nothing said about either moment generating functions or characteristic functions. The sections devoted to statistical inference are very weak. The likelihood ratio test is never mentioned. Maximum likelihood is inadequately discussed as is the whole Neyman-Pearson theory. Perhaps this criticism is due largely to the fact that Hald succeeds in accomplishing more than he originally set out to do, thereby suggesting that he could easily have accomplished still more.

An interesting and most welcome feature found at the conclusion of many of the chapters is a section devoted to "Notes and References." In this section, Hald describes the history of the contents of the chapter, and acknowl-

edges the authorship of many of the proofs of the theorems presented. For example, at the end of the chapter entitled "Fundamental Calculus of Probabilities," Hald writes:

> The calculus of probabilities was developed during the seventeenth century in connection with the solution of problems of games of chance. The classical definition of probability given in §1.4 took form in a correspondence between PASCAL and FERMAT in 1654. On basis of this definition an extensive mathematical theory was developed. Important contributions were made by J. BERNOULLI in *Ars Conjectandi*, 1713, and P. S. DE LAPLACE, who in *Théorie analytique des probabilités*, 1812, gave a very extensive and systematic exposition of the results and methods of the calculus of probabilities. . . .
>
> The construction of a calculus of probabilities from a set of axioms, as done in §1.3 is analogous to the way in which other branches of applied mathematics have been built up. The first systematic exposition of this method and its consequences was given by A. KOLMOGOROFF: *Foundations of the Theory of Probability*, Chelsea Publishing Company, New York, 1950, originally published in German in 1933. An elementary exposition of the modern mathematical concept of probability may be found in P. R. HALMOS: *The Foundations of Probability*, Amer. Math. Monthly, 51, 1944, 493–510. . . .
>
> An elementary exposition of the different concepts of probability from a philosophical point of view can be found in E. NAGEL: *Principles of the Theory of Probability*, International Encyclopedia of Unified Science, Vol. 1, No. 6, Chicago, 1939.

Published in conjunction with this book is a separate volume of statistical tables and formulas. This supplement is one of the most extensive books of tables published to date. Included in this volume are tables of the normal distribution, the t-distribution, the χ^2-distribution, the F-distribution, the distribution of the range, probits, confidence limits for the parameter of the binomial distribution, random numbers, etc.

Not only is *Statistical Theory* a worthwhile addition to the family of books in statistics, but it is unique among books on statistics in that a quotation from George Bernard Shaw is used to illustrate a point. Hald states:

> Thus the concepts of stochastic and causal dependence must be carefully differentiated. Bernard Shaw in his characteristic manner illuminates this point in the following quotation from the section "Statistical Illusions" of the preface to *The Doctor's Dilemma:* "Thus it is easy to prove that the wearing of tall hats and the carrying of umbrellas enlarges the chest, prolongs life, and confers comparative immunity from disease; for the statistics shew that the classes which use these articles are bigger, healthier, and live longer than the class which never dreams of possessing such things. It does not take much perspicacity to see that what really makes this difference is not the tall hat and the umbrella, but the wealth and nourishment of which they are evidence, and that a gold watch or membership of a club in Pall Mall might be proved in the same way to have the like sovereign virtues."

Modern Elementary Statistics. *John E. Freund.* New York: Prentice-Hall, Inc., 1952. Pp. x, 418. $5.50.

NORMAN RUDY, *Sacramento State College*

ATTEMPTS to shift the emphasis in introductory statistics courses from descriptive to inductive statistics have suffered from the absence of suitable textbooks and from the poor preparation in mathematics of most college students. The book reviewed here frankly emphasizes inductive statistics and, according to the preface and the publisher's advertisements, requires only a minimum amount of mathematics. On this basis alone, the book deserves the close attention of those concerned with the introductory statistics course.

The author states that "The order and the emphasis of the material covered follows the modern trend in the teaching of statistics—to include *informally* topics that in the past have often been taught only on an advanced level." At other points, he emphasizes that the meanings of statistical ideas are of greater importance than the formulas.

Despite certain criticisms to be made later, the reviewer's over-all impression is that the book will perform a valuable service in a general introductory course. The concept of a sampling distribution is very well developed and is emphasized whenever a statistic is introduced. Undoubtedly, the student who reads this book carefully will absorb what many believe to be the single most important idea in a beginning course, namely that of variability in the possible outcomes of a sample. The discussion of confidence intervals and their interpretation is also very good, and is strengthened by the construction of intervals for a non-normal distribution. The informal manner of presentation employed and the general tone of the book are consistent with what appears to be a current trend toward the humanizing of statistical ideas, comparable to the current humanism in the natural sciences.

The adoption of an informal manner of presentation, which is virtually a necessity in the introductory course, places upon the author the responsibility of providing enough exposition to insure that the meanings are clear. One of the weaknesses of the book lies in the fact that too often the meaning is subordinate to the formula, too often the elements of the formula are not justified heuristically or intuitively. A few examples will illustrate this criticism.

In the section entitled "Some Rules of Probability," the notions of conditional and joint probability are presented without benefit of any graphical or tabular aids, such as the two-way frequency table, found in the books by Wilks and Duncan. Again, in the discussion of confidence intervals for population means, using small samples, the factor $\sqrt{n-1}$ appears without any mention of degrees of freedom. In fact, the concept of degrees of freedom is not mentioned at any of the three places at which reference is made to the "t" distribution. In yet another instance, the χ^2 distribution is discussed in three different applications: first, as employed in the construction of confidence intervals for the population standard deviation; second, as a test for

association in contingency tables; and, third, as a test of goodness-of-fit. In the first application, the notion of degrees of freedom is not mentioned, while in the other two it is. There is no connection made between the two formulas given for χ^2, and no heuristic justification given either for squaring the difference between observed and expected frequencies or for dividing the square by the expected frequency.

The other general criticism which can be made of the book concerns the omission of some ideas of modern statistics which are important and basic enough to be considered in an introductory course. In the discussion of hypothesis testing the reader is advised to avoid the type II error (accepting the null hypothesis when it is false) by never accepting the null hypothesis at all. By this approach, the author precludes any possible discussion of the theory of statistical decision,[5] certainly as important a topic in the introductory course as the use of runs in testing for randomness. (The latter gets eleven pages of text and two pages of tables.) The omission is especially disturbing because in each of the two examples used, the decision to reserve judgment on the null hypothesis is also a decision whose consequences should be considered. With the type II error eliminated from consideration, the selection of the level of significance (only two-sided tests are considered) is made by following "the customary rule" of .05. This advice is not at all different in spirit from the 3σ rule found in older textbooks.

Another major omission, in the reviewer's judgment, is the complete absence from the sections on correlation and regression of the concept of a "model" or of any indication of the importance of *a priori* considerations in scientific investigations. According to the text, the treatment accorded the data (linear or curvilinear fit) is determined by inspection of the data itself, and since curvilinear fitting is taken as being outside the scope of the book, there is no need to discuss criteria for choosing between linear and curvilinear regressions. In the discussion of the standard error of estimate (in which no mention is made of degrees of freedom), it is incidentally revealed that X and Y are assumed to be normally distributed [a bivariate normal distribution?]. As usual, no mention is made of the fact that the prediction of Y on the basis of a given X is actually that of the average Y corresponding to the given value of X.

A final comment, lest these criticisms lead the reader to overlook or discount the comments made in the third paragraph: The reviewer has adopted the book for a general introductory course.

Elementary Statistics. Revised. *Morris Myers Blair* (Professor of Economics, University of Tulsa). New York: Henry Holt & Co., 1952. Pp. xiv, 735, $5.50.

R. Clay Sprowls, *University of California (Los Angeles)*

While reading the first chapter of this book, my pulse quickened in anticipation of a "new" book in the field of elementary statistics. The author states: "Because life is so short that we cannot know all about any-

thing, we have to make our decisions on the basis of sample information. This is the statistical method;" and, "Statistics as a science in its broader meanings is not only a means of guiding our judgment in making the daily decisions of life, but also a device or tool for discovering new truths. It is one of the most powerful engines for research." From here to the end of the book my pulse was normal except for one small flutter on page 208 where the author writes: "But analysis is usually not an end in itself. One analyzes in order that he may comprehend and forecast, infer, project, and apply his information for further practical or theoretical use. The great end of statistical description and analysis is *inference*." If the author really believes this, why did he not orient his textbook in this direction?

Is it because, as he says, "He is convinced that statistics should first be taught to most students from the standpoint of preparing them to be consumers of statistics instead of creators of research?" What better way to teach students to be intelligent consumers (even of descriptive statistics) than to make them create one small piece of research in which they themselves analyze a problem, specify the data needed, collect that data, make the necessary estimates, compute their errors, and then draw an inference?

Or is it because he thinks of a very selective area of consumption? "Many farmers and small merchants read carefully the price quotations, deliveries of grains and livestock . . . tables and charts . . . in the daily newspapers. These people are consuming statistics." I infer from this that he wishes to teach people how to interpret statistical data—an admirable wish. These people and others are also continually faced with making decisions which are based upon incomplete information and, therefore, subject to uncertain outcome. Implicitly or explicitly, they deal in probabilities. What better place than in elementary statistics to contribute to the general education by introducing the statistical method with its emphasis upon choice among alternatives and the probability of being wrong in that choice?

Granting that Blair is correct in appraising the consumer area, I believe that there are serious omissions in the text. I shall cite a few examples. In his discussion of the range (p. 152) he does not mention the interpretative fallacies inherent in the dependence of the magnitude of the range upon the sample size. In the discussion of regression analysis there is no cautioning about the value of the intercept often being an extrapolated value subject to large error and often an impossible value in the particular problem. Spurious correlation, a serious pitfall for the consumer, is not mentioned; neither is the frequently recurring trap of the regression fallacy. On p. 236 it is stated that rank correlation methods are not widely used. This *Journal* recently featured an article on the use of ranks which listed sixty-nine works in its bibliography; fifteen of these had the word rank in their title and others, I am sure, also discussed ranking methods. In the discussion of stratified sampling (p. 370), there is nothing to indicate that the real gains from stratification occur when the conditions are such that individual strata are sampled disproportionately to their size in the population. In the chapter on seasonal variation

(chapter 16) he gives three methods of calculating seasonal variation, but nowhere do I find (a) an explanation of a seasonal index of, say, 120 for a particular month; (b) any mention of adjusting data for seasonal variation; (c) an interpretation of a seasonally adjusted value.

It seems to me that Blair has not only missed a consumer area to which statistics has much to contribute, but that he has also failed to write a book which is successfully consistent with his own original intent, both because of these omissions which I have mentioned and because he continually drags the student through a mass of computations (a function of the producer?) to the neglect of interpretation. In my judgment, this is just another textbook in elementary statistics. It may be preferred to another because the first edition (1944) has been used in the past, because there are elaborate worksheets which aid students (a valuable service), or because it may be a convenient handbook. In other respects it is not distinguishable from a dozen or so other elementary textbooks in statistics.

Statistical Methods for Social Workers. *Wayne McMillen.* Chicago: The University of Chicago Press, 1952. Pp. xi, 564. $6.75.

DANIEL O. PRICE, *University of North Carolina*

THE stated aim of this book is to provide an orderly introduction to descriptive statistics for social workers (vii). In the opinion of this reviewer it misses the mark considerably. It is uneven in its treatment of material and contains several serious statistical flaws. It deals almost entirely with the field of descriptive statistics and contains the statement, "*Perhaps* in the long run inductive statistics will make contributions to human knowledge that clearly entitle it to be regarded as of far greater value to society than descriptive statistics" (vii, reviewer's italics). The first five chapters deal with the collection, editing, tabulation, and presentation of data.

Although the author seems dubious of the value of inductive statistics to social workers, the two illustrations on page 161 are both examples of inductive statistics.

In Chapter IX on the Mean, Median, and Mode it is not made clear that the "long" and "short" methods of computing the Mean produce identical results with the exception of rounding errors. In fact one gets the impression that the "short" method is an approximation to the results of the "long" method with the statement, "Thus the two methods produced substantially the same result" (p. 234).

Chapter X is a 17 page chapter on the Geometric Mean and Logarithms. Considering the doubtful value of the geometric mean to social workers this space might well have been utilized further in the book in the treatment of statistics of relationship.

In the chapter on Measures of Absolute Variability the author confuses the binomial and normal distributions, actually using a binomial distribution as his illustration on page 270 and calling it a normal curve. The binomial

distribution is not mentioned. In this same chapter eight pages are devoted to the mean deviation, a measure of limited usefulness, and only six pages are spent on the standard deviation, one of the most basic statistical measures. It is also in this same chapter that reference is made to Appendix P, a four page appendix on how to extract a square root. At no point is a reference made to *Barlow's Tables of Squares, Cubes, Square Roots, Cube Roots, and Reciprocals.*

Chapter XIII deals with Ratio Background or semilogarithmic paper. This material really belongs in the early chapter on graphic presentation.

Although the book deals primarily with descriptive statistics, Chapter XV is an 11 page chapter on Sampling which contains several errors sufficiently serious to deserve comment. No distinction is made between a systematic sample and a random sample, the formula for the estimated standard error of the mean does not include the loss of a degree of freedom for use of the sample standard deviation instead of the universe standard deviation, the t distribution is not even mentioned, and the probable error is used incorrectly. On page 340 we find the statement, "The mean of the sample plus and minus the probable error defines the range within which half the sample means may be expected to fall." Erroneous interpretations such as this are part of the reason for dropping the probable error from modern statistical usage. Confidence limits are not mentioned.

In Chapter XVI on Time Series a "derivation" of the two normal equations for getting the least squares straight line is given. The only difficulty is that what is given is *not* a derivation since the criterion of minimizing the sums of squares of deviations is not utilized and this characteristic of the least squares straight line is not mentioned until two pages later.

Chapter XVII on Correlation and Contingency leaves so much to be desired that by itself it would make the book of very limited value even if there were no shortcomings elsewhere in it. The *only* method given for computing a correlation coefficient is by first determining the regression line, then computing Y_c corresponding to each Y, finding the difference between Y and Y_c and squaring these differences, then finding the difference between each Y and the mean of Y and squaring these differences, then from these two sets of squares of differences determining the standard error of estimate and the standard deviation, and utilizing these two measures to find the correlation. The procedure in itself would discourage most students from ever computing a correlation coefficient. No real understanding of the correlation coefficient is shown and no test of statistical significance of a correlation is given. In fact the author says, "Unfortunately, this cannot be done" (p. 386). His confusion seems to lie in not distinguishing between statistical significance and practical importance.

The author gives two illustrations of high correlation between time series and suggests that the high relationship observed is due merely to chance, overlooking the fact that the correlation is due to the close relationship of each variable to a third variable, time.

As a suggestion for simplifying the computation of a correlation coefficient, the author suggests using the Spearman rank correlation as an estimate of the actual correlation between two variables and shows how this is computed. He gives no indication of the differences in assumptions and in utility of the two types of correlation coefficients.

As a whole the book has a point of view of statistics which was fairly general about 30 years ago. A tangible illustration of this is the fact that the book at no point assumes that the student has access to a calculating machine or even mentions a calculating machine. This is definitely not a modern statistics text and it is to be hoped that it will not set the tenor of statistical training among social workers.

Punched Cards: Their Applications to Science and Industry. *Robert S. Casey* and *James W. Perry*, editors. New York: Reinhold Publishing Corporation, 1951. Pp. viii, 506. $10.00.

HARRY P. HARTKEMEIER, *University of Missouri*

THIS book reveals the fact that chemists are now facing the same problem which confronted statisticians about 70 years ago. Statisticians trying to handle by hand the large volume of data collected in the 1880 U. S. Census took over seven years to organize and make available the information obtained. Chemists working now on the *Gmelin Handbuch der anorganischen Chemie* think that it will take 10 to 12 years to complete the work and they hope to complete the 8th edition "about 1960." This work is taking so long that many chemists fear, and others are convinced, that the era of the classical handbook is approaching its end. The parallel is striking. Statisticians in the '80's stated that the U. S. Census of 1890 would take more than 10 years to organize by hand and another census would have to be started before the information from the previous one could be made available, so some more rapid method of handling the data was imperative. This information was presented to some people who set about to invent and develop the Hollerith machines. Now, over half a century later, chemists are turning to the same punched-card machines to solve a similar problem—that of making available the pertinent information quickly, before it is made useless by the passage of time.

An alternative to periodic handbooks is the mechanical information center. Various parts of this book point out that similar problems have arisen in other fields, such as searching U. S. patent disclosures, hospital records, medical reports, library record cards, etc. Scientists have been forced to consider mechanical methods of searching indexes, and this book contains an excellent and clear discussion of the difficulties of using standard punched-card machines for searching such indexes. Although codes may be devised to classify books by size, date of publication, number of pages, etc., so that when a book is assigned to one class it is excluded from all others, a code for subjects presents a different problem, for the assignment of a book or article

to one subject class does not exclude it from all other subject classes. To classify a book under two or more subject classes requires the use of two or more vertical card fields. This also usually requires that the indexer decide which is the major subject field and which are minor subject fields. A person searching for all books which deal with a given subject might want to examine a book even though it is indexed under a minor field. As the number of fields increases, the standard machine cannot search all of them at once, or on one run of the cards, for all books dealing with a certain subject, or search for all books dealing with a combination of subjects.

A person, for example, who wants to locate all books dealing with the use of statistical methods to control the quality of a chemical compound used in a recently patented process to manufacture a new plastic would like to be able to locate all books dealing with this combination of subjects regardless of whether the book is primarily one on statistical analysis using this illustration of quality control methods, or primarily a book on plastics which includes a discussion of the fact that this particular plastic would not have been possible without uniform quality of a chemical compound obtained by using statistical methods of quality control, or even a book on patent laws that happened to contain an illustration of a court case arising out of the use of this plastic and involving the presentation of evidence including quality control charts. It would be nice if the machine would sort separately cards for books involving all of the subjects desired, all but one, all but two, etc. This book points out the need for machines that will accomplish such mechanical searches on one run of the cards and probably stimulated the research that resulted in the perfection of the machine described in *Library Applications of Punched Cards* by Ralph Parker (American Library Association, Chicago, 1952). The use of horizontal fields instead of vertical fields permits the machine to search 12 fields in succession on one run of the cards. Each of the 12 subject fields will accommodate a binary code number equivalent to 10 decimal places in the Dewey decimal system.

This book contains the case histories of many punched-card applications and the editors are to be congratulated upon the completion of a considerable project involving many people in widely scattered locations. "The hand-sorted edge-punched cards are discussed in greater detail than the machine-sorted cards. In fact, one object has been to make the book serve as an operating instruction manual for the edge-punched cards. It is not possible to do the same for machine-sorted cards within the scope of this book" (p. iii). The reviewer grants that this is the first and only book presenting detailed information on all hand-sorted edge-punched card systems, but he is also convinced that anyone adopting a hand-sorted system will find it to be only a temporary solution. Even for such prospective users the book may be disappointing because "there is no attempt to make critical comparisons" (p. 39). In a book written by so many people in different locations it is very difficult to avoid some duplication and a few statements that some may question. For example, "The only known correct way to sample is by the use of random

numbers" (p. 412), and "The mathematics of 'goodness of fit' have not yet been adequately investigated" (p. 413).

This book is not very suitable as a college textbook but it should be in most libraries to provide supplementary assignments. Many scientists, information specialists, executives, and office managers will find it useful as a reference.

Population, Food, and Economic Progress. *Merrill K. Bennett.* Houston, Texas: The Rice Institute, July 1952. Pp. 68. Paper.

COLIN CLARK, *The Bureau of Industry, Brisbane, Australia*

THIS pamphlet is a triumph of common sense. Such a world do we live in, that we are now beginning to regard common sense among learned men as one of the rarer virtues. A very well known professor of theoretical economics was recently heard to grumble that men of common sense never looked at statistics, and that men who studied statistics always seemed to leave their common sense behind them.

Dr. Bennett's work should now be well known, not only for his studies at the Food Research Institute of Stanford University, but also for his work in the field of international comparisons of levels of living, or "real income." (*American Economic Review*, 1951). We have not, and will not have for many years to come, accurate information about income levels in the under-developed countries, and the United Nations statistics of average income per head in these countries are not (speaking with all restraint) what they appear to be. Dr. Bennett devises a most ingenious statistical technique for combining all the scattered scraps of information available for most of these countries, such as the number of automobiles, telephones, high school students, and so on. The present reviewer had set out to measure the incomes of some of these countries quite differently, and the check with the Bennett method proved to be extremely satisfactory; so he is attempting to extend its sphere of application to more countries and to more recent years.

One of Dr. Bennett's principal figures is taken from FAO publications (which perhaps need no more qualification than they have received), namely calories of food consumption per head of population. He begins by reminding us of an obvious fact, the neglect of which has caused endless trouble—namely, that if statistics show that Europeans and North Americans consume 3,000 calories per head per day, and the inhabitants of southeast Asia 2,000, the latter figure, *per kilogram of average body weight*, may be greater than the former.

The next point which he makes is that the greater part of the world did and does obtain its calories from grains and roots, but that we all like to consume a substantial proportion of animal foodstuffs when we can afford them. In the past, this was the privilege of the wealthy on the one hand, of nomadic huntsmen and herders on the other; for obvious reasons, the latter category

now represents a far smaller proportion of the earth's population than it did in the past.

The next obvious fact which we tend to forget, and to which Dr. Bennett forcefully draws our attention, is that throughout the greater part of human history mankind must have had an inappreciable, if any, rate of population increase. In 1650, when modern population estimates begin, the world had a population of about five hundred millions. Biologists have contended that the human race has been in existence for five hundred thousand years, though some of them have recently raised the estimate to a million years—even if we assume that they have a couple of zeros out of place, we are still left with the same problem. Whether we assume that the human race started with two people, as Christians believe, or take any larger number predicated for us by biologists, and whatever conceivable date we take as the starting point of the human race, we are still left with the conclusion that any rate of population increase even remotely comparable with those which we know now would have produced a far larger world population by 1650. This accords with the observations of anthropologists that primitive people produce large families, but that there is little or no net rate of population increase among them. Dr. Bennett has employed a team of historical students who follow world population data back to the year 1,000 A.D., and find plenty of evidence of stagnant or indeed declining populations, in all those historical periods when political order broke down. There is some evidence of heavy population declines, both in Europe and Asia, during the centuries of warfare and disorder between 300 and 800 A.D.

The reviewer would, however, disagree with Dr. Bennett's results on India. There, it appears, population rose to a maximum about the middle of the seventeenth century, under the despotic but ordered rule of the Mogul emperors; and in the ensuing two centuries of anarchy and warfare remained stationary or declined until order was again restored under British rule in the nineteenth century.

The writings of Malthus, which have enjoyed such a comeback among American intellectuals in the present generation, are thus found to be a theoretical speculation not valid historically—nor, for that matter, did Malthus make a valid estimate of the possibilities of agricultural development. Human capacity to reproduce is not limited. Demographic studies of primitive peoples, showing an average of six or seven children born to a woman who lives to the end of her reproductive period, agree well with the estimates of the medical sub-commission of the recent British royal commission on population, as to what would be the total average fertility of the modern English woman, if she married early and imposed no restrictions on reproduction; and also agree with recorded total fertility in some modern communities like Brazil, and sample areas in China. But under primitive conditions, or under the conditions of warfare and anarchy with which so much of the history of civilization is disfigured, this rate of reproduction will hardly allow the human race to maintain its numbers. Population increases only

occur where mankind can establish a stable political order—a surprisingly difficult thing to do—and at the same time, accumulate sufficient medical knowledge to control the ravages of disease—a combination only found in the modern world, except perhaps for some periods under the ancient Greeks and Egyptians.

And when such order and increase have been established, it is surprising how short is the time interval—as both modern experience and ancient history teach us—before mankind, of its own accord, starts in various ways to reduce its rate of reproduction. We know our own nineteenth century history; in Japan and Russia we can trace, with reasonable accuracy, the fall in the average total fertility from six to four over the short period of the last thirty or forty years; and very few people are aware that a similar trend is occurring in India, where the rate of population increase is rapidly decelerating.

In his final and magnificently common sense conclusion, therefore, Dr. Bennett sees "a touch of hysteria" in "current attitudes towards what is called the population problem," or the fashionable idea that America ought to send what might be called missionaries of contraception to inform the benighted Asiatics. Even in the modern world, areas of excessive population density are rare, and most of the habitable world is still uninhabited. War and anarchy in the past have not arisen in over-populated regions—rather the reverse, among the people of sparsely settled grazing areas, who wished to live on the labor of others, rather than cultivate land which was available for settlement. Calculations of the remote future at which the world, at its present rate of increase, will become finally overcrowded, are of little interest —a purely arithmetic exercise, and about as sterile as the favorite old computation of what would happen if a man left one thousand dollars to be invested at compound interest for the benefit of his descendant in six hundred years time.

Census of Manufactures: 1947: Indexes of Production. *Bureau of Census and Board of Governors of the Federal Reserve System.* U. S. Government Printing Office, 1952. Pp. viii, 99. $1.75.

PAUL B. SIMPSON, *University of Oregon*

THIS volume explains the material, methods, and results of computation of of an index of manufacturing based on the 1947 and 1939 Census of Manufactures. It extends the work of Solomon Fabricant: *The Output of Manufacturing Industries*, 1899–1937 (National Bureau of Economic Research, New York, 1940), both in years covered and methods employed. The bulk of the work comprises detailed tables of quantity and value of output, industry indexes of output, employment, and weights. The text explains methods used, and the reasons therefor.

The final index for 1947 relative to 1939 is 174, indicating that this eight-year period witnessed a record percentage rate of manufacturing growth in the United States. The percentage increase from 1921 to 1929 was somewhat

more, namely 88 as against 74, but this was accomplished during passage from a depressed year, 1921, to a peak year, 1929, whereas 1939 was a near record year of industrial production. Jim Corbett fans may be interested to know that in the ten-year period 1899 to 1909, manufacturing increased 56 per cent. The growth since pre-war years has been large by nearly any standard.

The greatest task facing any index computor is obtaining data comparable over time. The Census and Federal Reserve Board collaborators have done a remarkable job in obtaining as much such data as they have. The work involved in such an undertaking may be illustrated with a random quotation:

> Census figures for "other natural cheese" in 1947 represent "shipments and interplant transfers." These figures are not comparable with 1939 "production" statistics, since they exclude natural cheese made into process cheese. . . . Hence a United States Department of Agriculture quantity figure, which includes this production was used for 1947" (p. 77).

It requires a special talent of patience, imagination, and devotion to accuracy to ferret out such detailed information. The statistical world and the nation owes a debt to those who have repeated the task a thousand times over in preparing this index and other economic series. The quotation also serves to remind us that pretentious economic categories such as supply, marginal efficiency of capital, and gross national product are compoundings of matters no more mysterious than cheese.

The Census Bureau could confine its work to collecting information and turning it over to the public for analysis. The present volume is ample testimony that further stages of analysis are desirable. The comparison of 1947 and 1939 data was accomplished better on home grounds than it could have been elsewhere. Indeed, this reviewer believes that there would be a net gain if the scope of the index construction activity had been expanded still further. Measurement cannot be separated from economic analysis, and the closest integration is desirable. Below are some examples of points where the computation of the index raises problems of general economic analysis.

The index formula used is a cross-weight formula, the Marshall-Edgeworth formula using 1939 and 1947 value-added-per-unit weights. With simple 1939 weights, the index would have been 184, and with 1947 weights, the index would have been 169. This difference indicates that quantity of production increased most for those items whose prices increased least. In supply and demand terms, this says that for those items whose demand schedules increased the most, supply schedules increased proportionately more. If we look not at the total manufacturing index but at industry groups, a related tendency is manifest. The difference between group indexes based on the two weights is greatest where production increased the most, the 1947 weights yielding a lower index. Thus leather products, whose physical production increased only 15 per cent from 1939 to 1947, had commensurate price increases, with the result that the two weight systems yielded the same index. On the other hand non-electrical machinery increased in production

177 per cent from 1939 to 1947 as measured by the 1939 weight index, and by 160 per cent as measured by the 1947 weight index. The Spearman coefficient of rank correlation of 1939 indexes and the difference of 1947 and 1939 weight indexes for nineteen groups is .74. Assume that for those groups where growth was greatest, the dispersion of the amounts of growths of individual industries is largest. This assumption, similar to one frequently made in stratified sampling, is reasonable here. Using this assumption, we reach the conclusion that within most groups of industries, as in the total manufacturing group, supply reacted proportionately more strongly to larger increases in demand.

Let us suppose provisionally that the explanation of these facts lies in economies of large scale production. Such economies would presume downward sloping long-run supply curves, and would become operative exactly in those industries whose capital expansion was the greatest, and those would likely be industries where demands increased most. The observed facts would be explained nicely. We now have a new understanding of the different meanings of the indexes based on 1939 and 1947 weights respectively. The 1939 weight index approximates the number of years it would have taken the 1939 industry, with its cost structure, to produce the 1947 output. The 1947 weight index is the reciprocal of the fraction of a year that it would have taken the 1947 industry to produce the 1939 output. The cross weight index is a measure based on the operation of some intermediately adjusted productive system. The problem of weighting methods has found solution in terms of economic meaning.

It may be objected that the suggested explanation may not be the correct one. There are indeed other possible explanations of the response of supply to demand. Monopoly and wage rate effects are possibilities. Another is the observable fact that two groups of industries showing less than average growth in volume of output, namely petroleum and coal products and lumber and lumber products, are closely tied to natural resources. Were such resource limitations the cause of the low output expansion and high price reactions? To what extent do measures of manufacturing production reflect such limitations? A discussion of the questions rising from theory and measurement would be very illuminating.

Another aspect of the weighting problem is its relation to national income measures. Value added by manufacture is the basis for desirable weighting, because it eliminates duplication in value of commodities. However, from the standpoint of the economy as a whole, duplication may still arise, because taxes, advertising, insurance, professional services and the like are represented in value added by manufacture, though they have separate representation in national income. The United Nations Statistical Office has recommended the use of weights excluding such economic activities. The writers of the work in review have handled the problem only by noting that preparation of such weights, others than those based on value added by manufacture, "was not found feasible" (p. 3). Surely the question of the position of the manufactures index in all output deserves more consideration than this.

It is worth noting in this connection that the differences of national income and value added by manufacture may be just as important as changes in value added by manufacture per unit of output in time. Total value added by manufacture in 1947 was $74,428 billions, of which 12.2 per cent or $9,053 billion was due to food production. National income originating in manufacturing, according to Department of Commerce estimates, was $59,459 billions in 1947 of which 9.8 per cent, or $5,822 billions, were allocated to food products. The differences demonstrate that this weighting question warrants careful analysis. Also it would be desirable to compare the production indexes with other value and price information than Census data, which information might have been used for independent testing of the consistency of different source data.

The definition of physical output as value of output at constant prices would work well if prices stayed constant, if the physical nature of goods remained unchanged, and if the definition is significant to economic analysis. All three, however, are questionable. Quality changes particularly embarrass the authors. They observe: "Technological change may make possible the substitution of less expensive materials without affecting the quality of the product" (p. 7), and again, "The (quality) changes are difficult if not impossible to measure quantitatively, and the failure to reflect them in the present index, results in a downward bias" (p. 7). These two statements imply that a product is defined as a capacity for doing a particular job in production functions and consumer preference functions. If this is to be our definition of output, let us have it in the open and apply it the best we may. (It will lead to surprising results in many cases such as rubber tires, and nylon socks.) It also appears at conflict with a weight selection which reflects cost conditions in different industries. Economic theory has various definitions of output. These should not be ignored.

Questions of economic analysis arise in connection with the otherwise excellent studies of representation of the "missing" industries, that is, industries for which physical production data of any sort are largely missing, covering about one-fourth of all industries in terms of value added by manufacture. The selection of a method for measuring output in these industries was based on studies of what method would have worked best in industries where quantity data are available, if they had not been available. A critical test was to determine which of three artificially constructed indexes came closest to the originally computed index. The assumptions used in constructing these artificial indexes were: (1) temporal change in value of output per unit of physical product was similar among related industries, (2) temporal change in value added by manufacture per unit of physical product was similar among related industries, and (3) the temporal change in output per man employed was similar among industries. The winner in this interesting contest was number three, which yielded the highest index for 1947 by some six points, or three per cent in the case of the output assumption (1), and by half as much for the value added assumption (2).

Economic theory poses some questions about this selection of method. The "missing" industries are particularly prominent in machinery industries, fabricated metal products, and miscellaneous industries, which are industries whose production has grown particularly fast. This fact suggests a question of the following kind: Is the price behavior of these industries to be explained by inefficiencies arising from dynamic change, by increasing efficiencies due to economies of scale, or by wage rate developments? So far as we can judge by price changes, the answer would appear to be that the second possibility is most likely, since generally price rises were less in the more expanding industries. This suggests, in turn, that output per man rose more in the "missing" industries than in others, suggesting, in turn, that the output per man selection may have understated the index, though by less than the other two alternatives considered. A discussion of these theoretical questions would be illuminating both to economic analysis and to measurement.

The reviewer believes that theory or explanation is inseparable from measurement. At the moment, measurement is ahead of theory in the sense that we have measures such as the index of manufacturing whose significance is not fully understood because we do not know the underlying forces accounting for the particular observed changes. Measurement would be improved if answers were sought to specific questions of the following nature:

(1) How long would it have taken the productive mechanism of 1939 (say) to produce the output of 1947, and conversely with interchange of dates?

(2) To what extent was the change in output accomplished by changes in employment, changes in technology, and economies of scale?

(3) What was the output of consumers' goods, in the sense of selling values at constant prices with allowance for quality changes and in the sense of production requirements? What was the investment output in the sense of production requirements and in the sense of capacity for new production?

(4) What are the indexes of output and employment in national income and gross national product senses?

(5) How well do the indexes reconcile with price, value of shipment, national income and other independent data?

These questions cannot be answered easily or quickly, but beginnings can be made at once. Perhaps a committee of learned societies is in order for the purpose of making recommendations regarding expanded activities of government statistical workers.

Washington State Statistical Abstract. *Marilyn Druck Robinson.* Seattle, Washington: University of Washington Press, 1952. Pp. xi, 159. $4.50. Paper.

PAUL B. SIMPSON, *University of Oregon*

THIS volume brings together a large quantity of data, covering principally area, population, employment, payrolls, production, trade, transportation, income, prices, and finance. The state of Washington nat-

urally receives the most attention, but comparative data for the United States and for western states are frequently presented.

Although this volume uses data gleaned from census and other national sources as its chief stock in trade, it is far from a mere compendium of such material. The reviewer estimates that at least one-fourth of the 104 tables contain data not available from standard sources. Examples of information hard to come by are trade and income statistics based on Washington tax information, indexes of business activity in the Northwest, housing- and rental-survey information, family income-survey information, Washington data about utilities and motor carriers, and lumber shipments reported by the Pacific Lumber Inspection Bureau. Other tables give inter-census interpolations and county and regional break-downs of special information not available in standard sources. Payrolls, births and deaths, population, labor force, income, and the mineral statistics are examples.

The work is carefully documented as to sources, nature of information, and additional sources, and includes warnings about misusing the data. By careful study the researcher can obtain a fairly complete picture of what regional information is available from all sources. There can never be too much explanation, however. In using the volume this reviewer found himself wishing that it included a summary of the idiosyncrasies of the Washington State retail and industrial income tax laws and of covered employment under the Social Security Program. A separate list of Washington State reports and local source materials would also have been useful to give the research worker a handy reference to sources of local data.

Very few similar volumes exist for other states. The Universities of Alabama and of Mississippi have compiled statistical abstracts for their states, and a certain amount of statistical information is included in the state blue books. Generally, regional data are difficult to locate and to obtain. Marilyn Robinson and the other members of the Bureau of Business Research of the University of Washington who prepared the abstract have made a valuable contribution to regional research.

PUBLICATIONS RECEIVED

Baer, Reinhold. *Linear Algebra and Projective Geometry.* New York: Academic Press, Inc., 1952. $6.50.

Beyer, Glenn H., and Yantis, Theodore R. *Practices and Precepts of Marketing Prefabricated Houses.* Washington, D. C.: Housing and Home Finance Agency, November 1952. Paper. 35 cents from U. S. Government Printing Office.

Bliss, C. I. *The Statistics of Bioassay.* New York: Academic Press, Inc., 1952. $3.50.

Bolza, Hans. *Grundriss einer systematischen Wirtschaftslehre, I Band.* Stuttgart: W. Kohlhammer, 1947. Paper.

Chapman, Agatha L. *Wages and Salaries in the United Kingdom, 1920–38.* New York: Cambridge University Press, 1953. $9.50.

Coombs, Clyde H. *A Theory of Psychological Scaling.* Ann Arbor: University of Michigan Press, 1952. Paper. $1.75.

Copeland, Morris A. *A Study of Money Flows in the United States.* New York: National Bureau of Economic Research, 1952. $7.50.

Council of Europe, Secretariat-General of. *A Few Aspects of Population and Economy: Statistical Study of Certain Aspects of the Structure of Population and of the Economy of the Member States of the Council of Europe.* Strasbourg, 1952. Paper.

Darwin, Charles Galton. *The Next Million Years.* Garden City, N. Y.: Doubleday and Company, 1953.

Dublin, Louis I. *Publications and Addresses, 1903–1952.* New York: Metropolitan Life Insurance Company, 1952. Paper.

Eddy-Rucker-Nickels Company. *Progress in Productivity and Pay: All U. S. Manufacturing, 1914–1947.* Cambridge, Mass., 1952. Paper. $2.50.

Fabricant, Solomon. *The Trend of Government Activity in the United States Since 1900.* New York: National Bureau of Economic Research, 1953. $4.00.

Garrett, Henry E. *Statistics in Psychology and Education. Fourth Edition.* New York: Longmans, Green and Co., 1953. $5.00.

Gore, W. L. *Statistical Methods for Chemical Experimentation.* New York: Interscience Publishers, Inc., 1952. $3.50.

Grampp, William D., and Weiler, Emanuel T., eds. *Economic Policy, Readings in Political Economy.* Homewood, Ill.: Richard D. Irwin, Inc., 1953. $3.75.

Gruenberger, Fred. *Computing Manual.* Madison: University of Wisconsin Press, 1952. Paper.

Halcrow, Harold G. *Agricultural Policy of the United States.* New York: Prentice-Hall, Inc., 1953. $5.50.

Hartree, D. R. *Numerical Analysis.* London: Oxford University Press, 1952.

Heady, Earl O. *Economics of Agricultural Production and Resource Use.* New York: Prentice-Hall, Inc., 1952. $9.75.

Henderson, John S. *Production and Consumption.* University: University of Alabama Press, 1952. Paper. $2.00.

Institute for Economic Research, Osaka, Japan. *Causes of Decline in the World's Cotton Textile Trade.* Osaka: Toyo Spinning Co., Ltd., 1952. Paper.

International Labour Office. *Year Book of Labour Statistics, 1951–52. Twelfth Issue.* Geneva: 1952. Paper. $5.00; cloth, $6.00.

International Tin Study Group. *Statistical Year Book 1952—Tin: Tinplate: Canning.* The Hague: 1952. $5.60.

Kuznets, Simon, and Goldsmith, Raymond. *Income and Wealth of the United States. Income and Wealth Series II.* Cambridge: Bowes and Bowes, 1952. 35s.

Lane, Frederic C., and Riemersma, Jelle C., eds. *Enterprise and Secular Change, Readings in Economic History.* Homewood, Ill.: Richard D. Irwin, Inc., 1953. $5.00.

Lewis, Edward E. *Methods of Statistical Analysis in Economics and Business.* Boston: Houghton Mifflin Co., 1953. $5.50.

Iowa State College, Statistical Laboratory. *The WOI-TV Audience—A Report of a Bench Mark Survey of Households within 50 miles of Ames.* Ames, Iowa: 1952. Paper.

Machlup, Fritz. *The Economics of Sellers' Competition.* Baltimore: The Johns Hopkins Press, 1952. $6.50.

Machlup, Fritz. *The Political Economy of Monopoly.* Baltimore: The Johns Hopkins Press, 1952. $5.50.

Milne, William Edmund. *Numerical Solution of Differential Equations.* New York: John Wiley and Sons, 1953. $6.50.

National Bureau of Economic Research. *Conference on Research in Income and Wealth. Studies in Income and Wealth, Volume Fifteen.* New York: 1953. $3.50.

Neyman, Jerzy. *Lectures and Conferences on Mathematical Statistics and Probability. Second Edition, Revised and Enlarged.* Washington, D. C.: Graduate School, Department of Agriculture, 1952.

Okayama University, Mathematical Journal of. Vol. 1, Nos. 1–2, March 1952. Okayama, Japan: Department of Mathematics, Faculty of Science.

Organisation of European Economic Co-

operation. *Measurement of Productivity.* Paris: 1952. Paper. $1.25.

Quenouille, M. H. *Associated Measurements.* New York: Academic Press, Inc., 1952.

Schlote, Werner. *British Overseas Trade from 1700 to the 1920's.* Translated by W. H. Chaloner and W. O. Henderson. Oxford: Basil Blackwell, 1952. 22s 6d.

Stockton, John R., Henshaw, Richard C., Jr., and Graves, Richard W. *Economics of Natural Gas in Texas.* Austin: University of Texas, 1952. $5.00.

Tinbergen, J. *On the Theory of Economic Policy.* Amsterdam: North-Holland Publishing Co., 1952. Paper. $1.80.

Tippett, L. H. C. *The Methods of Statistics. Fourth Edition.* New York: John Wiley and Sons, 1952. $6.00.

Tolles, N. Arnold, and Raimon, Robert L. *Sources of Wage Information: Employer Associations.* Ithaca, N. Y.: New York State School of Industrial and Labor Relations, Cornell University, 1952. Paper. $3.00.

Ullmann, John E. *A Guide to Sales Quota Setting.* New York: Chemonomics, Inc., 1952. Paper. $3.00.

United Nations, Food and Agriculture Organisation of the. *Fruit and Vegetables: Production, Trade and Policies in Europe, 1947–1951.* Rome, Italy: 1952. Paper. $0.50.

U. S. Bureau of Foreign and Domestic Commerce, Office of Business Economics. *Foreign Aid by the United States Government, 1940–1951.* Washington, D. C.: U. S. Government Printing Office, 1952. Paper. $1.00.

U. S. Bureau of Human Nutrition and Home Economics. *Collection Methods in Dietary Surveys.* Southern Cooperative Series Bulletin 23. Washington, D. C.: U. S. Government Printing Office, 1952. Paper.

U. S. Bureau of Old-Age and Survivors Insurance. *Handbook of Old-Age and Survivors Insurance Statistics: Employment, Wages, and Insurance Status of Workers in Covered Employment, 1949.* Washington, D. C.: U. S. Government Printing Office, 1952. Paper. $1.00.

Vickrey, William S. *The Revision of the Rapid Transit Fare Structure of the City of New York.* New York: Mayor's Committee on Management Survey of the City, 1952. Paper.

Whelpton, P. K., and Kiser, Clyde V., eds. *Social and Psychological Factors Affecting Fertility, Volume 3.* New York: Milbank Memorial Fund, 1952. Paper. $1.00.

Williams, John H. *Economic Stability in a Changing World, Essays in Economic Theory and Policy.* New York: Oxford University Press, 1952.

Wold, Herman, and Juréen, Lars. *Demand Analysis.* New York: John Wiley and Sons, 1952. $7.00.

RANDOM DIGITS (6501–6875)

From *A Million Random Digits*, to be published by The Rand Corporation, Santa Monica, California

Digits 6001–6500 were published in Vol. 48, p. 167 (March 1953)

14322	91065	07841	36130	05313
61079	96910	72906	07361	47511
81115	49625	48799	89485	25072
88559	92476	32903	68009	60329
63075	13270	84758	49560	38527
81488	17340	74154	42801	27917
51667	37589	87147	24743	48023
99004	70322	60832	76636	56907
68656	66492	35933	52293	47953
38074	74083	09337	07965	65047
35428	01020	80680	59328	08712
49919	86379	74508	33579	77114
15092	48498	09938	44420	13484
82639	41800	95363	54142	17482
90326	63026	96712	79883	39225

1953 OFFICERS AMERICAN STATISTICAL ASSOCIATION

President	*President-Elect*	*Past President*
WILLIAM G. COCHRAN	HERBERT MARSHALL	ARYNESS JOY WICKENS
	Vice Presidents	
MORRIS H. HANSEN	WALTER E. HOADLEY, JR.	RENSIS LIKERT
	Directors	
W. J. DIXON	HOWARD L. JONES	JOHN W. TUKEY
MARGARET J. HAGOOD	WILLIAM R. LEONARD	RALPH J. WATKINS
	Secretary-Treasurer	
	SAMUEL WEISS	
	Members of the Council	
LUCILE DERRICK	HOWARD LEVENE	HARRY S. SCHWARTZ
A. ROSS ECKLER	WILLIAM G. MADOW	JOHN R. STOCKTON
MARTIN R. GAINSBRUGH	KENNETH E. MILLER	RUFUS S. TUCKER
EZRA GLASER	ALMARIN PHILLIPS	DAVID F. VOTAW, JR
WERNER HOCHWALD	ORVILLE B. RAILEY	F. M. WADLEY
A. J. JAFFE	J. A. RIGNEY	W. ALLEN WALLIS
MILDRED M. LAUDER		W. J. YOUDEN

A cumulative Index to Volumes 1–34, 1888–1939, may be obtained from the Office of the Secretary of the American Statistical Association.
Reprints of all articles published since 1945 may be purchased from the office of the secretary. Articles which have appeared during the past twelve months are usually available in quantity.